

THE CAMBRIDGE
ENCYCLOPEDIA OF THE
**World's Ancient
Languages**

Edited by

ROGER D. WOODARD

Andrew V. V. Raymond Professor of the Classics

Professor of Linguistics

University of Buffalo (The State University of New York)



CAMBRIDGE
UNIVERSITY PRESS

LP. O. 73

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK

40 West 20th Street, New York, NY 10011-4211, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

Ruiz de Alarcón 13, 28014 Madrid, Spain

Dock House, The Waterfront, Cape Town 8001, South Africa

<http://www.cambridge.org>

© Cambridge University Press 2004

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2004

Reprinted 2004

Printed in the United Kingdom at the University Press, Cambridge

Typeset Minion 10/12 pt and Formata *System* L^AT_EX 2_ε [TB]

A catalogue record for this book is available from the British Library

ISBN 0 521 56256 2 hardback

Dedicated to the memory of Henry Hoenigswald
Scholar, gentleman, friend

Abbreviations

Any abbreviation that deviates from the form given below is noted within the text of the individual chapter or within a chapter-specific list.

Linguistic terms

abl.	ablative
abs.	absolutive
acc.	accusative
act.	active
adj.	adjective
adv.	adverb (adverbial)
all.	allative
anim.	animate
aor.	aorist
art.	article
asp.	aspirated
aux.	auxiliary (verb)
caus.	causative
cl.	clause
coll.	collective
com.	common
comp.	comparative
comt.	comitative
conj.	conjunction
conjv.	conjunctive
conn.	connective
cons.	consonant
constr.	construct (state)
cont.	continuant
cop.	copula
dat.	dative
def. art.	definite article
dem.	demonstrative
det.	determinate
detv.	determinative
dial.	dialect

dir.	directive
dir. obj.	direct object
disj.	disjunctive
du.	dual
dur.	durative
emph.-pcl.	emphatic particle
encl.	enclitic
eq.	equative
erg.	ergative
ext.	extended
fem.	feminine
final-pcl.	final-particle
fut.	future
gdve.	gerundive
gen.	genitive
ger.	gerund
impf.	imperfect
impftv.	imperfective
impv.	imperative
inan.	inanimate
inc.	inclusive
indef. art.	indefinite article
indet.	indeterminate
indic.	indicative
inf.	infinitive
instr.	instrumental
interr.	interrogative
intr.	intransitive
iter.	iterative
juss.	jussive
loc.	locative
mediopass.	mediopassive
mid.	middle

Abbreviations

Any abbreviation that deviates from the form given below is noted within the text of the individual chapter or within a chapter-specific list.

Linguistic terms

abl.	ablative	dir.	directive
abs.	absolutive	dir. obj.	direct object
acc.	accusative	disj.	disjunctive
act.	active	du.	dual
adj.	adjective	dur.	durative
adv.	adverb (adverbial)	emph.-pcl.	emphatic particle
all.	allative	encl.	enclitic
anim.	animate	eq.	equative
aor.	aorist	erg.	ergative
art.	article	ext.	extended
asp.	aspirated	fem.	feminine
aux.	auxiliary (verb)	final-pcl.	final-particle
caus.	causative	fut.	future
cl.	clause	gdve.	gerundive
coll.	collective	gen.	genitive
com.	common	ger.	gerund
comp.	comparative	impf.	imperfect
comt.	comitative	impftv.	imperfective
conj.	conjunction	impv.	imperative
conjv.	conjunctive	inan.	inanimate
conn.	connective	inc.	inclusive
cons.	consonant	indef. art.	indefinite article
constr.	construct (state)	indet.	indeterminate
cont.	continuant	indic.	indicative
cop.	copula	inf.	infinitive
dat.	dative	instr.	instrumental
def. art.	definite article	interr.	interrogative
dem.	demonstrative	intr.	intransitive
det.	determinate	iter.	iterative
detv.	determinative	juss.	jussive
dial.	dialect	loc.	locative
		mediopass.	mediopassive
		mid.	middle

N.	noun
neg.	negative
neut.	neuter
nom.	nominative
NP	noun phrase
num.	number
obj.	object
obl.	oblique
opt.	optative
part.	participle
pass.	passive
pcl.	particle
per.	person
perf.	perfect
perfv.	perfective
perfvz.	perfectivizer
pert.	pertinentive
pl.	plural
pluperf.	pluperfect
poss. suff.	possessive suffix
postp.	postposition
PP	prepositional phrase
prec.	precative
preC.	preconsonantal
pref.	prefix
prep.	preposition
pres.	present
pret.	preterite
preV.	prevocalic
pro.	pronoun
prosp.	prospective
quot.	quotative particle
refl.	reflexive
rel. pro.	relative (pronoun)
rel./connec.	relative/connective
sg.	singular
soc.	sociative case
SOV	Subject–Object–Verb (word order)
spec.	specifier
stat.	stative
subj.	subject
subjunc.	subjunctive
subord.	subordinate/subordinator/ subordination marker
subord.-pcl.	subordinating particle
suff.	suffix
splv.	superlative
s.v.	<i>sub voce</i>

top.	topicalizer
tr.	transitive
V.	verb
var.	variant
vent.	ventive
voc.	vocative
vow.	vowel
VP	verbal phrase

Languages

Akk.	Akkadian
Ar.	Arabic
Ass.	Assyrian
Av.	Avestan
Bab.	Babylonian
Cis. Gaul.	Cisalpine Gaulish
Eg.	Egyptian (Old, Late, Earlier)
Eng.	English
Etr.	Etruscan
Gk.	Greek
Gmc.	Germanic
Go.	Gothic
Hisp.-Celt.	Hispano-Celtic
Hitt.	Hittite
IE	Indo-European
Lat.	Latin
Lep.	Lepontic
Luv.	Luvian
Lyc.	Lycian
MA	Middle Assyrian
MB	Middle Babylonian
NA	Neo-Assyrian
NB	Neo-Babylonian
OA	Old Assyrian
O. Akk.	Old Akkadian
O. Av.	Old Avestan
OB	Old Babylonian
OHG	Old High German
OP	Old Persian
PG	Proto-Greek
PGmc.	Proto-Germanic
PIE	Proto-Indo-European
PIIr.	Proto-Indo-Iranian
PIr.	Proto-Iranian
PMS	Proto-Mije-Sokean
PS	Proto-Semitic
PSo.	Proto-Sokean
SB	Standard Babylonian

Skt. Sanskrit
Sum. Sumerian
Y. Av. Young Avestan

Other

abbr. abbreviation

dict. dictionary
intro. introduction
lit. literally
NA not applicable
NS new series
trad. traditional
translit. transliteration

N.	noun
neg.	negative
neut.	neuter
nom.	nominative
NP	noun phrase
num.	number
obj.	object
obl.	oblique
opt.	optative
part.	participle
pass.	passive
pcl.	particle
per.	person
perf.	perfect
perfv.	perfective
perfvz.	perfectivizer
pert.	pertinentive
pl.	plural
pluperf.	pluperfect
poss. suff.	possessive suffix
postp.	postposition
PP	prepositional phrase
prec.	precative
preC.	preconsonantal
pref.	prefix
prep.	preposition
pres.	present
pret.	preterite
preV.	prevocalic
pro.	pronoun
prosp.	prospective
quot.	quotative particle
refl.	reflexive
rel. pro.	relative (pronoun)
rel./connec.	relative/connective
sg.	singular
soc.	sociative case
SOV	Subject–Object–Verb (word order)
spec.	specifier
stat.	stative
subj.	subject
subjunc.	subjunctive
subord.	subordinate/subordinator/ subordination marker
subord.-pcl.	subordinating particle
suff.	suffix
splv.	superlative
s.v.	<i>sub voce</i>

top.	topicalizer
tr.	transitive
V.	verb
var.	variant
vent.	ventive
voc.	vocative
vow.	vowel
VP	verbal phrase

Languages

Akk.	Akkadian
Ar.	Arabic
Ass.	Assyrian
Av.	Avestan
Bab.	Babylonian
Cis. Gaul.	Cisalpine Gaulish
Eg.	Egyptian (Old, Late, Earlier)
Eng.	English
Etr.	Etruscan
Gk.	Greek
Gmc.	Germanic
Go.	Gothic
Hisp.-Celt.	Hispano-Celtic
Hitt.	Hittite
IE	Indo-European
Lat.	Latin
Lep.	Lepontic
Luv.	Luvian
Lyc.	Lycian
MA	Middle Assyrian
MB	Middle Babylonian
NA	Neo-Assyrian
NB	Neo-Babylonian
OA	Old Assyrian
O. Akk.	Old Akkadian
O. Av.	Old Avestan
OB	Old Babylonian
OHG	Old High German
OP	Old Persian
PG	Proto-Greek
PGmc.	Proto-Germanic
PIE	Proto-Indo-European
PIIr.	Proto-Indo-Iranian
PIr.	Proto-Iranian
PMS	Proto-Mije-Sokean
PS	Proto-Semitic
PSo.	Proto-Sokean
SB	Standard Babylonian

Introduction

ROGER D. WOODARD

1.1 Preliminary remarks

What makes a language ancient? The term conjures up images, often romantic, of archeologists feverishly copying hieroglyphs by torchlight in a freshly discovered burial chamber; of philologists dangling over a precipice in some remote corner of the earth, taking impressions of an inscription carved in a cliff-face; of a solitary scholar working far into the night, puzzling out some ancient secret, long forgotten by humankind, from a brittle-leaved manuscript or patina-encrusted tablet. The allure is undeniable, and the literary and film worlds have made full use of it.

An ancient language is indeed a thing of wonder – but so is every other language, all remarkable systems of conveying thoughts and ideas across time and space. And ancient languages, as far back as the very earliest attested, operate just like those to which the linguist has more immediate access, all with the same familiar elements – phonological, morphological, syntactic – and no perceptible vestiges of Neanderthal oddities. If there was a time when human language was characterized by features and strategies fundamentally unlike those we presently know, it was a time prior to the development of any attested or reconstructed language of antiquity. Perhaps, then, what makes an ancient language different is our awareness that it has outlived those for whom it was an intimate element of the psyche, not so unlike those rays of light now reaching our eyes that were emitted by their long-extinguished source when dinosaurs still roamed across the earth (or earlier) – both phantasms of energy flying to our senses from distant sources, long gone out.

That being said, and rightly enough, we must return to the question of what counts as an ancient language. As *ancient* the editor chose the upward delimitation of the fifth century AD. This *terminus ante quem* is one which is admittedly “traditional”; the fifth is the century of the fall of the western Roman Empire (AD 476), a benchmark which has been commonly (though certainly not unanimously) identified as marking the end of the historical period of *antiquity*. Any such chronological demarcation is of necessity arbitrary – far too arbitrary – as linguists accustomed to making such diachronic distinctions as *Old English*, *Middle English*, *Modern English* or *Old Hittite*, *Middle Hittite*, *Neo-Hittite* are keenly aware. Linguistic divisions of this sort are commonly based upon significant political events and clearly perceptible cultural shifts rather than upon language phenomena (though they are surely not without linguistic import as every historical linguist knows). The choice of the boundary in the present concern – the ancient-language boundary – is, likewise (as has already been confessed), not mandated by linguistic features and characteristics of the languages concerned.

However, this arbitrary choice, establishing a *terminus ante quem* of the fifth century, is somewhat buttressed by quite pragmatic linguistic considerations (themselves consequent

to the whim of historical accident), namely the co-occurrence of a watershed in language documentation. Several early languages first make a significant appearance in the historical record in the fourth/fifth century: thus, Gothic (fourth century; see Ch. 36), Ge'ez (fourth/fifth century; see Ch. 14, §1.3.1), Classical Armenian (fifth century; see Ch. 38), Early Old Georgian (fifth century; see Ch. 40). What newly comes into clear light in the sixth century is a bit more meager – Tocharian and perhaps the very earliest Old Kannada and Old Telegu from the end of the century. Moreover, the dating of these languages to the sixth century cannot be made precisely (not to suggest this is an especially unusual state of affairs) and it is equally possible that the earliest attestation of all three should be dated to the seventh century. Beginning with the seventh century the pace of language attestation begins to accelerate, with languages documented such as Old English, Old Khmer, and Classical Arabic (though a few earlier inscriptions preserving a “transitional” form of Arabic are known; see Ch. 16, §1.1.1). The ensuing centuries bring an avalanche of medieval European languages and their Asian contemporaries into view. Aside from the matter of a culturally dependent analytic scheme of historical periodization, there are thus considerations of language history that motivate the upper boundary of the fifth century.

On the other hand, identifying a *terminus post quem* for the inclusion of a language in the present volume was a completely straightforward and noncontroversial procedure. The low boundary is determined by the appearance of writing in human society, a graphic means for recording human speech. A system of writing appears to have been first developed by the Sumerians of southern Mesopotamia in the late fourth millennium BC (see Ch. 2, §§1.2; 2). Not much later (beginning in about 3100 BC), a people of ancient Iran began to record their still undeciphered language of Proto-Elamite on clay tablets (see Ch. 3, §2.1). From roughly the same period, the Egyptian hieroglyphic writing system emerges in the historical record (see Ch. 7, §2). Hence, Sumerian and Egyptian are the earliest attested, understood languages and, *ipso facto*, the earliest languages treated in this volume (on the problem of undeciphered languages like Proto-Elamite, see §1.2).

It is conjectured that humans have been speaking and understanding language for at least 100,000 years. If in the great gulf of time which separates the advent of language and the appearance of Sumerian, Proto-Elamite, and Egyptian societies, there were any people giving written expression to their spoken language, all evidence of such records and the language or languages they record has fallen victim to the decay of time. Or the evidence has at least eluded the archeologists.

Though no human language is documented prior to the late fourth millennium BC, it is still possible in certain instances to recover the linguistic system of a deeply archaic, preliterate people, using the remarkable methodology commonly dubbed the *comparative method of historical linguistics* (see Ch. 45). The development of the comparative method and the discovery of the linguistic principles which make it possible was one of the greatest, if perhaps one of the less recognized, scientific achievements of the nineteenth century. Following upon the pioneering efforts of philologists such as the English jurist, Sir William Jones, scholars like Rasmus Rask, Franz Bopp, August Schleicher, Jacob Grimm, Karl Verner, Karl Brugmann, and Hermann Osthoff, among still others (on whom, see, *inter alia*, Mallory 1989:9–23, Lehmann 1967), developed the comparative method and applied it in the reconstruction of Proto-Indo-European, the parent language of the Indo-European language family. Though spoken between the fifth and third millennia BC and nowhere attested by written record, the grammar and lexicon of the language are well known through their reconstruction (see Ch. 17).

The comparative method has likewise been used to reconstruct the parent of the Semitic languages, Proto-Semitic (see Ch. 6, §§2–3), spoken sometime prior to the third

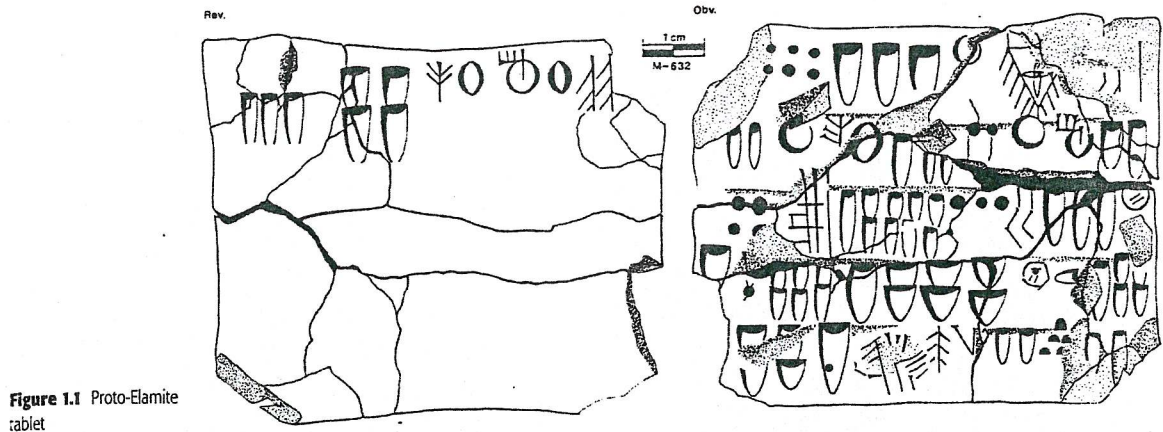


Figure 1.1 Proto-Elamite tablet

millennium BC. The reconstruction of Proto-Semitic's own parent language, Proto-Afro-Asiatic, from which are also descended Egyptian, Berber, Cushitic, Chadic, and perhaps still other languages is ongoing (see Ch. 6, §1).

With the exception of only a few languages of Mesoamerica, the native American languages of antiquity are known solely through reconstruction. Those exceptions are provided by the Mayan languages (see Ch. 43, §1) and Epi-Olmec (see Ch. 44, §§1–2); though the Zapotec language is perhaps attested as early as about the sixth/fifth century BC, the inscriptional evidence is very meager and the language is poorly understood at present (see below, §1.3.8).

Not every language which is attested in the period that extends from the beginning of written records to the fifth century AD is treated in this volume. There are generally two reasons for this exclusion: (i) the written remains of the language can be read (to a greater or lesser extent), but the evidence of the language which is provided by these records is sufficiently meager to limit significantly a knowledge of the language and, consequently, to proscribe any attempt to offer a meaningful grammatical description of it; and (ii) the written remains of the language cannot be understood – in other words, the recorded language has not yet been deciphered. Languages falling into the latter category will be addressed first, and then those of the former.

1.2 Undeciphered languages

1.2.1 Undeciphered Elamite

The earliest attested undeciphered script (late fourth millennium BC) is the one called *Proto-Elamite*. Tablets bearing this script have been recovered in large numbers from the same Iranian region in which ancient Elamite is attested, hence the name Proto-Elamite (see Fig. 1.1). Far fewer in number are the inscriptions from the same general area which are written in the script called *Linear Elamite*. For discussion of each of these, see Chapter 3, §§2.1–2.2 (see also Englund 1996).

1.2.2 The Indus Valley language

In the middle of the third millennium BC, writing emerges in the archeological record of the Harappan culture of the Indus Valley. The characters of this Indus Valley script (see Fig. 1.2) are of a well-developed, somewhat conventionalized pictographic nature at the earliest

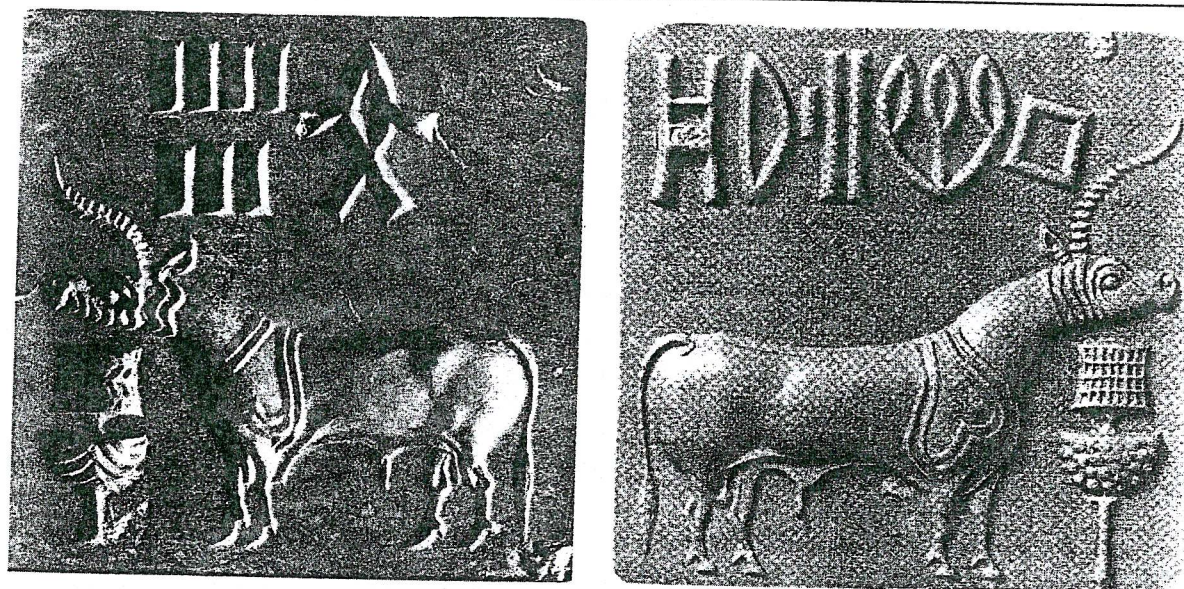


Figure 1.2 Indus Valley inscriptions

phase of the script's attestation (possibly suggesting some earlier unattested developmental stage). The number of characters identified likely reveals that the script operates with both logograms (symbols representing entire words) and syllabograms (phonetic symbols having the value of a syllable). Lying behind the Indus Valley script may well be a Dravidian language (see Ch. 42, §1) or possibly an early form of Indo-Aryan (see Ch. 26, §1). On the Indus Valley script and its attempted decipherment, see Parpola 1996 and 1994.

1.2.3 Cretan languages

Examples of three distinct, undeciphered scripts have survived in the remains of the Minoan civilization of ancient Crete. The oldest of these is called *Cretan Hieroglyphic* or *Cretan Pictographic* (see Fig. 1.3) and its use is dated to the period 2000–1600 BC, seal stones providing the bulk of examples. The pictographic symbols making up the script probably have a syllabic value.

The second of the undeciphered Cretan scripts is known from only a single document, the *Phaistos Disk* (dated to about 1700 BC; see Fig. 1.4). The disk has been the object of repeated attempts at decipherment since its discovery in the early twentieth century. While success has often been claimed, none of the proposed decipherments carries conviction.

Linear A, the third of the Minoan scripts, is the best represented of the three. Dating from about the mid-nineteenth to mid-fifteenth centuries BC, Linear A documents partially overlap chronologically with those written in Cretan Hieroglyphic, though in terms of historical development, the former may trace its origins to the latter. Linear A, in turn, appears to be the source of the Mycenaean Greek script, Linear B (see Ch. 25, §§1.1; 1.2; 2.1), though a simple direct linear descent is not probable. Of the three Minoan scripts, Linear A holds the greatest hope for decipherment. Recent work by Brown (1990) and Finkelberg (1990–1991) has taken up a notion proposed by Palmer in the middle of the twentieth century (e.g., Palmer 1968) which would identify the Linear A language as a member of the Anatolian subfamily of Indo-European. On the Cretan scripts see, *inter alia*, Chadwick 1990; Palaima 1988; Woodard 1997.



Figure 1.3 Cretan
Hieroglyphic inscription
and portrait stamped on
a sealing



Figure 1.4 The Phaistos
Disk (side A)

Mention should also be made of the undeciphered language called *Eteo-Cretan*. Much later than the three Bronze Age Minoan scripts, Eteo-Cretan is preserved in inscriptions written in the Greek alphabet. On Eteo-Cretan, see Duhoux 1982.

1.2.4 Cypriot languages

Prior to the emergence of Greek writing on Cyprus, attested by about the middle of the eleventh century BC (and the somewhat later appearance of Phoenician; see Ch. 11, §1.2; Ch. 24, §2), the island was inhabited by a people, or by groups of people, who were recording their speech in the undeciphered set of scripts called *Cypro-Minoan* (see Table 1.1). As the name suggests, these Cypriot writing systems appear to have their origin in a writing system of Minoan Crete, Linear A being the likely candidate. *Archaic Cypro-Minoan* is the name given to the script found on only a single inscription, dated to about 1500 BC.

Table 1.1 A partial inventory of Cypro-Minoan characters

I	𐀀	𐀁	𐀂	𐀃	𐀄
𐀅	𐀆	𐀇	𐀈	𐀉	𐀊
𐀋	𐀌	𐀍	𐀎	𐀏	𐀐
𐀑	𐀒	𐀓	𐀔	𐀕	𐀖
𐀗	𐀘	𐀙	𐀚	𐀛	𐀜
𐀝	𐀞	𐀟	𐀠	𐀡	𐀢
𐀣	𐀤	𐀥	𐀦	𐀧	𐀨
𐀩	𐀪	𐀫	𐀬	𐀭	𐀮
𐀯	𐀰	𐀱	𐀲	𐀳	𐀴

















































This script has been analyzed as the likely ancestor of the more widely attested *Cypro-Minoan 1*, found in use between approximately the late sixteenth and twelfth centuries BC. A distinct script, *Cypro-Minoan 2*, has been found on thirteenth-century documents from the site of Enkomi. Yet a third, *Cypro-Minoan 3*, dating also to the thirteenth century BC, has turned up not on Cyprus but in the remains of the ancient Syrian city of Ugarit (see Ch. 9, §1; on the Cypro-Minoan scripts, see especially Masson 1974, 1977; Palaima 1989).

Cypro-Minoan 1 appears to have provided the graphic model for the Greek syllabary of Cyprus (see Ch. 25, §2.2). This Greek syllabic script was in turn not only used for writing Greek but also adopted for some other language of Cyprus, as yet undeciphered, dubbed *Eteo-Cypriot*. The Eteo-Cypriot inscriptions are commonly regarded as the documentary remains of an indigenous people of Cyprus who had withstood assimilation to the communities of Greek and Phoenician settlers. After Greek and Phoenician settlement of Cyprus, Eteo-Cypriots appear to have concentrated particularly in the area of Amathus (on the Eteo-Cypriot inscriptions, see Masson 1983:85–87).

1.2.5 Byblic language

As well as the various Canaanite and Aramaic scripts and languages preserved in the archaeological remains of Syria-Palestine that are treated in this volume (see Chs. 6, 9–13), there is an additional script, attested by a small number of inscriptions, which is commonly (though not universally) regarded as undeciphered. In the course of his excavations at the site of the ancient city of Byblos (Biblical Gebal) on the coast of the modern state of Lebanon, the French archeologist Maurice Dunand unearthed inscriptions, on bronze and stone, executed in a previously unknown script. Many of the symbols are of a hieroglyphic nature, some apparently descended from or inspired by characters of the Egyptian hieroglyphic script; the Byblian script thus bears the tag *Pseudo-Hieroglyphic*, or, less commonly, *Proto-Byblic*. The script, judging by the number of identified symbols (114 by Dunand's analysis), is likely syllabic. As early as 1946 (a year after Dunand's publication of the inscriptions), the decipherment of Byblian Pseudo-Hieroglyphic was announced by a distinguished French philologist, Edouard Dhorme, who read the language of the script as Phoenician. Dhorme's proposed decipherment and others which have followed (see Daniels 1996:29–30 for discussion of subsequent attempts) have not been received with confidence and the script and its language still reside in the undeciphered column.

Table 1.2 Characters of the Meroitic script

Character			Character		
Hieroglyphic	Cursive	Transcription	Hieroglyphic	Cursive	Transcription
		a			l
		e			h
		i			h
		o			s
		y			se
		w			k
		b			q
		p			t
		m			te
		n			to
		ne			d
		r			word-divider

1.2.6 North African languages

In the ancient Nubian city of Meroë (in the north of modern Sudan – the great city which Herodotus calls the μητρόπολις τῶν ἄλλων Αἰθιοπίων, “the capital of all Ethiopia”), the Egyptian scripts must have been long known and utilized. By the third century BC, however, with the rise of the Meroitic kingdom, a native writing system appeared and continued in use for recording Meroitic language until the fourth century AD. Two varieties of the script are known: a hieroglyphic script based on Egyptian Hieroglyphic, and a cursive form based on Egyptian Demotic (see Ch. 7, §2.1).

The phonetic values of the symbols of the Meroitic writing system, unlike those of many of the undeciphered languages discussed thus far, have been purportedly identified. The majority of symbols have been assigned the value of a single consonant or vowel sound (i.e., the script is analyzed as fundamentally alphabetic), with a small set of syllabic CV (consonant + vowel) symbols filling out the inventory of characters (compare Ugaritic’s consonantal script, supplemented by three CV characters; see Ch. 9, §2.2); see Table 1.2. While Meroitic texts can thus be given a phonetic reading, the language uttered in such a reading cannot be understood with the exception of a very few words, chiefly proper nouns. On the Meroitic script and language, see Wenig 1982, Griffith 1911, 1912.

For the *Ancient Libyan* or *Numidian* script, or scripts, which have been read as recording archaic Berber, but which some would regard as undeciphered, see Chapter 6, §1.1.3 with references. On the partially deciphered *Proto-Sinaitic* inscriptions, see Chapter 12, §2.2.

1.2.7 European languages

From Portugal and Spain come ancient inscriptions recorded in those scripts called *Iberian*, broadly divided into two groups, Northeast and South Iberian. The latter group includes the variety of the script called *Turdetan*, after the ancient Turdetanians, of whom the Greek geographer Strabo wrote: “These are counted the wisest people among the Iberians; they write with an alphabet and possess prose works and poetry of ancient heritage, and laws

Table 1.3 Irish Ogham (Craobh-Ruadh); font courtesy of Michael Everson

Symbol	Transcription	Name	Symbol	Transcription	Name
┐	b	beithe	┌	h	úath
└	l	luis	┘	d	dair
┘	f	fern	┐	t	tinne
┘	s	sail	┘	c	coll
┘	n	nin	┘	q	ceirt
┘	m	muin	┘	a	ailm
┘	g	gort	┘	o	onn
┘	ng	géal	┘	u	úr
┘	z	straif	┘	e	edad
┘	r	ruis	┘	i	idad
✕	ea	ébad	◊	oi	ór
✕	ia	iphín	┘	ui	uilen
┘	ae	emancholl			

composed in meter, six thousand years old, so they say" (*Geography* 3.1.6). One form of the Northeast Iberian writing system was adopted by speakers of Celtic for recording their own language (*Hispano-Celtic* or *Celtiberian*; see Ch. 35, especially §2.1), and these Celtic documents are interpretable (for the language, see Ch. 35, especially §§3.1; 3.4; 4.2.1.1; 4.3.6; 5.1). However, the Iberian scripts were used principally for a language or languages which are not understood, in spite of the fact that there also occur Iberian-language (*Old Hispanic*) inscriptions written with the Greek and Roman alphabets, and even bilingual texts. On the Iberian scripts and language(s) see, *inter alia*, Untermann 1975, 1980, 1990, 1997; Swiggers 1996; Diringer 1968:193–195.

While the South Picene language of eastern coastal Italy appears to be demonstrably Indo-European (belonging to the Sabellian branch of Italic; see Ch. 33), the genetic affiliation of its meagerly attested northern neighbor, North Picene, remains uncertain (though the two were formerly lumped together under the name *East Italic* or *Old Sabellian*). Though completely readable (being written in an Etruscan-based alphabet), North Picene remains largely impenetrable, in spite of the fact that a Latin-North Picene bilingual exists (a brief inscription, the identity of the non-Latin portion of which has been disputed). For an examination toward a tentative translation of the long North Picene inscription, the *Novilara Stele*, see Poultney 1979 (providing a summary of earlier attempts at interpretation).

The documentation of Insular Celtic – the Celtic languages of Ireland and Britain – (as opposed to Continental Celtic; see Ch. 35) which has survived from antiquity is very meager indeed, and is limited to Irish. The script used in recording this early Irish is the unusual alphabetic system called Ogham (see Table 1.3); most of its characters consist of slashing lines, longer and shorter (notches being used at times for vowel characters), giving the impression that it was originally designed to be “written” by means of an ax or some similar sharp instrument, with wood serving as a medium. The Ogham inscriptions, which date as early as the fourth century AD (and perhaps as early as the second century), can be read (owing to our knowledge of later Irish) but consist largely of personal names and provide little data on which can be constructed a linguistic description of Ogham Irish. For such descriptions of Insular Celtic, the linguist must await the appearance of Old Irish and Old

Welsh manuscripts in about the eighth century AD (and hence Ogham Irish is not treated in the present volume).

There is, however, a second ancient language of Britain which is written with a variety of Ogham, the language of *Pictish*. The Picts, who receive their name from Latin *Picti* "painted ones" (presumably referring to the practice of tattooing, though other etymologies have been proposed), inhabited portions of modern Scotland, along with the Scots, a Celtic people of Irish origin. A much broader, earlier distribution of the Picts has also been claimed. The Picts are known for their production of stone monuments on which are engraved intriguing images of animals and other designs, at times accompanied by Ogham inscriptions. The language of the Pictish Ogham inscriptions is not understood; it is not Celtic and probably not Indo-European. On the Pictish language, see Jackson 1980; for Ogham generally, see McMannus 1991.

1.3 Insufficiently attested languages

The differences between the languages of this group and the preceding are in some cases only a matter of degree (rather than one of kind), and not sharply one of intelligible versus unintelligible.

Among those recorded languages of antiquity which can be read and understood to an appreciable degree but which were judged too meagerly attested to be included in the present volume of grammatical descriptions, several are languages which were spoken on the Italian peninsula and the neighboring island of Sicily.

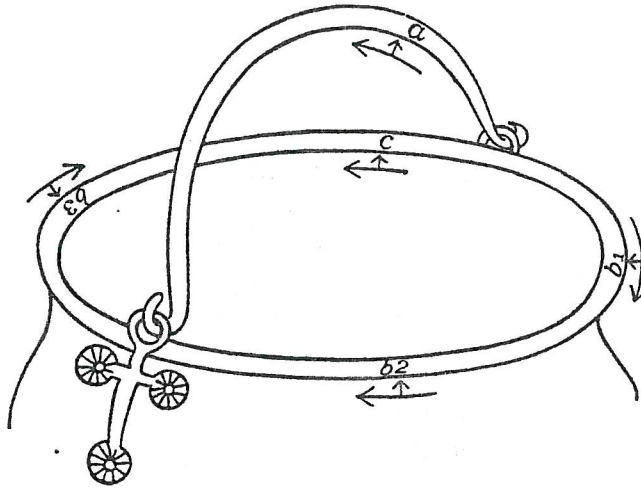
1.3.1 Sicel

From Sicily come several inscriptions written in a language which appears to be Indo-European; a number of glosses are claimed as well (see Conway, Whatmough, and Johnson 1933 II:449–458; on Sicel generally, see Pulgram 1978:71–73 with references). The name assigned to the language, Sicel or Siculan, is that given by Greek colonists to the native peoples of Sicily whom they there encountered in the eighth century BC. Little is known about the ethnicity of these Siceli. The form *esti* occurs in Sicel, seemingly the archetypal Indo-European "(s)he is." Interpretations of other inscriptional forms show considerable variation. Tradition held that the Siceli had migrated to Sicily from the Italian peninsula: thus, Varro (*On the Latin Language* 5.101) writes that they came from Rome; Diodorus Siculus (*Library of History* 5.6.3–4) records that the Siceli had come from Italy and settled in the region of Sicily formerly occupied by a people called the Sicani. On the basis of the available linguistic evidence, however, Sicel cannot be demonstrated to be a member of the Italic subfamily of Indo-European (see Ch. 32, §1).

On the inscriptional fragments from western Sicily identified as *Elymian*, see Cowgill and Mayrhofer 1986:58 with references.

1.3.2 Raetic and Lemnian

From the eastern Alps, homeland of the tribes called Raeti by the Romans, come a very few inscriptions in a language which has been claimed to bear certain Indo-European characteristics. For example, from an inscription carved on a bronze pot (the Caslir Situla; see Fig. 1.5) comes the Raetic form *-talina* which has been compared to Latin *tollo* "I raise" (see Pulgram 1978:40 with additional references). However, similarities to Etruscan have



α lavisešeli

β 1, 2 velxanu | lup·nu pitiave |

3 kusenkustrinaxe

γ φelna vinutalina.

Figure 1.5 The Casir Situla

also been identified and the two are perhaps to be placed in a single language family, along with a language attested on the island of Lemnos in the north of the Aegean Sea. Lemnian is known principally from a single inscribed stele bearing the engraved image of a warrior, dated to the sixth century BC. On these connections, see Chapter 39, §1.

Of the Raeti, the Roman historian Livy (*History* 5.33.11) writes, following upon his discussion of the Etruscans: “Undoubtedly the Alpine tribes also have the same origin, particularly the Raeti, who have been made wild by the very place where they live, preserving nothing of their ancient ways except their language – and not even it without corruptions.”

1.3.3 Ligurian

The Ligurians were an ancient people of northwestern Italy. Writing in the second century BC, the Greek historian Polybius (*Histories* 2.16.1–2) situates the Ligurians on the slopes of the Apennines, extending from the Alpine junction above Marseilles around to Pisa on the seaward slopes and to Arezzo on the inland side. Another Greek, Diodorus Siculus (*Library of History* 5.39.1–8), writes of the Ligurians eking out a life of hardship in their heavily forested, rock-strewn, snow-covered homeland and of the extraordinary stamina and strength which this lifestyle engendered in both men and women.

The Ligurian language appears to be attested in certain place names and glosses, some of which have been assigned Indo-European etymologies. For example, Pliny the Elder, a Roman author of the first century AD, in describing the grain called *secale* in Latin, noted

that its Ligurian name (the name among the Taurini) is *asia* (*Natural History* 18.141). If the Ligurian form was once *sasia* (see Conway, Whatmough, and Johnson 1933 II:158), then, it has been proposed, the word may find relatives in Celtic – Welsh *haid* and Breton *heiz* “barley.” The location of its speakers, abutting Celtic areas (and Strabo writes of Celtoligurians; *Geography* 4.6.3), might itself be taken to suggest an affiliation with the Indo-European family, but such a relationship cannot be confirmed by the available linguistic evidence.

1.3.4 Illyrian

The historical peoples called Illyrian occupied a broad area of the northwest Balkans. Evidence for an Indo-European intrusion into the region can be identified by the late third millennium BC; an identifiable “Illyrian” culture appears only in the Iron Age (see, *inter alia*, Wilkes 1992:28–66). By the first century AD, the Greek geographer Strabo, in describing that part of Europe south of the Ister (the Danube) can identify as Illyrian those people inhabiting the region bounded on the east by the meandering Ister, on the west by the Adriatic Sea, and lying above ancient Epirus (*Geography* 7.5.1). For the Romans, the province of *Illyricum* denotes a rather larger administrative area. The term “Illyrian” can, however, be used by classical authors to designate a variety of peoples in and beyond the Balkans (see the discussion in Katičić 1976:156–163).

Within the northwestern Balkan region itself there was considerable cultural diversity, with not only the so-called Illyrian tribes being present, but Celts as well, by at least the third century BC. Strabo writes of the Iapodes dwelling near Mount Odra (close to the border of modern Slovenia and Croatia) whom he calls a mixed Celtic and Illyrian tribe (*Geography* 4.6.10) and who, he adds, use Celtic armor but are tattooed like the Illyrians and Thracians (*Geography* 7.5.4; on the Thracians see §1.3.5). In his account of the wars which various Illyrian tribes waged against one another and against the Romans, the Greek historian and Roman citizen, Appian of Alexandria, writing in the second century AD, preserves a tradition in which one hears echoes of such Balkan ethnic diversity. Appian (*Roman History* 10.2) records that the Illyrians received their name from Illyrius, a son of Polyphemus (the cyclops of Homer’s *Odyssey*) and the nymph Galatea, and that Illyrius has two brothers, Celtus and Galas, namesakes of the Celts and the Galatae (the latter commonly being synonymous with “Celt” and perhaps used here to invoke descent from Galatea).

The Illyrian language presents an unusual case. While the Illyrians are a well-documented people of antiquity, not a single verifiable inscription has survived written in the Illyrian language (on two proposed Illyrian inscriptions, one demonstrably Byzantine Greek, see Katičić 1976:169–170). Even so, much linguistic attention (perhaps a disproportionately large amount) has been paid to the language of the Illyrians. Chiefly on the basis of Illyrian place and personal names, the language is commonly identified as Indo-European. To provide but two examples, the frequently attested name *Vescleves* has been etymologized as a reflex of Proto-Indo-European **wesu-klewes* (“good fame”), with Sanskrit *Vasuśravas* being drawn into the analysis; the place name *Birziminium*, interpreted as meaning “hillock,” has been traced to the Proto-Indo-European root **b^herǵ^h-*, source of, *inter alia*, Germanic forms such as Old English *beorg* “hill” (see Katičić 1976:172–176 for discussion). This onomastic evidence is supplemented by the survival of just a very few glosses of Illyrian words; for example, the Illyrian word for “mist” is cited as *rhinos* (ῥινός) in one of the scholia on Homer; see Katičić 1976:170–171, who compares Albanian *re*, earlier *ren*, “cloud.” Extensive study of Illyrian was undertaken by Hans Krahe in the middle decades of the twentieth

century, who, along with other scholars, argued for a broad distribution of Illyrian peoples considerably beyond the Balkans (see, for example, Krahe 1940); though in his later work, Krahe curbed his view of the extent of Illyrian settlement (see, for example, Krahe 1955). Radoslav Katičić (1976:179–180) has argued, on the basis of a careful study of the onomastic evidence, that the core onomastic area of Illyrian proper is to be located in the south-east of that Balkan region traditionally associated with the Illyrians (centered in modern Albania).

The modern Albanian language, it has been conjectured, is descended directly from ancient Illyrian. Albanian is not attested until the fifteenth century AD and in its historical development has been influenced heavily by Latin, Greek, Turkish, and Slavic languages, so much so that it was quite late in being identified as an Indo-European language. Its possible affiliation with the scantily attested Illyrian, though not unreasonable on historical and linguistic grounds, can be considered little more than conjecture barring the discovery of additional Illyrian evidence.

1.3.5 Thracian

At the northern end of the Aegean Sea, stretching upward to the Danube, lived in antiquity people speaking the Indo-European language of Thracian. The ancestors of the Iron Age Thracians had probably arrived in the Balkans as a part of the movement which brought the forebears of the Illyrians. For the Greeks, Thrace was a place wild and uncultivated, home to both savage Ares and Dionysus, god of wine who inspired frenzy and brutality in his worshipers. Herodotus (*Histories* 5.3; 9.119) writes of the Thracian practices of human sacrifice and widow immolation, and of the enormous population of the Thracians (second only to the Indians) and their lack of political unity. Were they unified, surmises the historian, they would be the most powerful people on the face of the earth.

Though the Thracian language is not well preserved, its attestation, unlike that of Illyrian, is sufficient to place its membership in the Indo-European family practically beyond doubt. A few short Thracian inscriptions survive (see Brixhe and Panayotou 1994a:185–188), but more valuable are the numerous glosses (e.g., *bólinthos* “European bison,” cf. Old Norse *boli* “bull”; *brûtos* “beer,” cf. Old English *breowan* “to brew”) coupled with the evidence of place and personal names. For a summary of the evidence see Katičić 1976:138–142; Brixhe and Panayotou 1994a:188–189; see also Cowgill and Mayrhofer 1986:54–55, with references. Onomastic evidence may suggest the occurrence of a language boundary within the Thracian area, demarcated by Mount Haemus. South of this boundary the language evidenced has been distinguished as Thracian, while that to the north has been called Daco-Mysian.

According to Greek tradition, the Phrygians of Anatolia had migrated from the Balkans (see Herodotus, *Histories* 7.73, who writes that the Phrygians were formerly called the Briges and had been neighbors of the Macedonians; on the Macedonians see §1.3.6), a view with which modern scholarship is generally in agreement. The Phrygian language does show certain similarities to Thracian, and some linguists have argued for linking the two in a single linguistic unit (Thraco-Phrygian). The appropriateness of the subgrouping is, however, uncertain; see Chapter 31, §1.5.

1.3.6 Macedonian

North of the Greeks, bracketed by Illyrians and Thracians, lived the Macedonians. Much uncertainty surrounds the linguistic status of the Macedonian peoples. Though, under the patronage of Macedonian kings, Philip the Second and his son Alexander the Great, Greek

culture would be spread across the Mediterranean and Near Eastern world and the Greek language would become a lingua franca (the Attic-based Koine dialect; see Ch. 24, §1) spoken from Italy to India, it remains unclear if Greek was the native language of the Macedonians (see Brixhe and Panayotou 1994b:206–207 for a synopsis of ideas about the identity of Macedonian).

To be sure, the Greek orator Demosthenes, in the fourth century BC, can revile and lambaste Philip as one of the *barbaroi* (“barbarians,” those who do not speak Greek, i.e., those who *babble*; *Orations* 3.17) and rehearse how in the old days the Macedonian king had been rightly subject to the Greeks, as *barbaroi* should be (*Orations* 3.24). He can skewer Philip with the charge that, not only is he not a Greek and unrelated to the Greeks, he is not even a *barbaros* from some worthwhile place, but he is a plague out of Macedonia – a place from which you cannot even acquire a good slave (*Orations* 9.31). A century earlier, Herodotus had told the story of an ancestor of Philip, Alexander the First (a contemporary of Herodotus), who had been allowed to compete in games at Olympia – though *barbaroi* were excluded from the competition – because he was able to demonstrate satisfactorily that he himself was descended from a Greek banished from Argos (*Histories* 5.22; 8.137–139).

Explicit references to “Macedonian speech” exist. Plutarch, the Greek savant of the first and second centuries AD, when writing of Cleopatra (*Life of Antony* 27.3–4), the last of the Ptolemies (the Macedonian kings of Egypt), lauds her linguistic abilities, reporting that she could speak the languages of the Ethiopians, Troglodytes, Hebrews, Arabs, Syrians, Medes, and Parthians. In contrast, her male predecessors had not even learned Egyptian and some had even “ceased to speak Macedonian” (μακεδονίζειν ἐκλιπόντων). Presumably they had continued to speak Greek (i.e., had not taken a vow of silence). Athenaeus, a Greek writer of the later second century AD, in his account of a “Learned Banquet” (*Deipnosophistae* 3.121f–122a), places on the lips of one of the guests, the cynic Cynulcus, a Latin word *decocta* (a kind of drink made by boiling and then rapidly cooling a liquid); in turn, Athenaeus has another guest, Ulpian (an “Atticist,” promoting the use of untainted Attic Greek), rebuke Cynulcus for uttering a barbarism (!). Cynulcus fires back, retorting that even in the best old Greek one finds Persian loanwords and that he knows many Attic Greeks “using Macedonian speech” (μακεδονίζοντας; a participle from Plutarch’s verb). Elsewhere, Plutarch uses an adverb *makedonistí* (μακεδονιστί) having the same sense. For example, in his *Life of Alexander* (51.4), Plutarch recounts how the Macedonian conqueror, in a fit of rage, refusing to be quieted by his body guards, shouted out for the *hypaspistai* (Macedonian infantry troops, one contingent of the army of Alexander), “calling in Macedonian – and this was a sign of a great disturbance.” The precise sense of “speaking Macedonian” in these and other passages can be and has been debated; yet when these references to Macedonian speech are considered in their context, it is not difficult for one to conclude that what is being reported is the use of a distinct, non-Greek (“barbarian”) Macedonian language.

In contrast, however, other classical authors explicitly identify the Macedonians as a Greek people. Polybius, the Greek historian of the second century BC, for example, describes Macedonians and Greeks as being *homophylos* (ὁμόφυλος) “of the same race” or “akin” (*Histories* 9.37.7). For references to other, similar texts, see Katičić 1976:107–108.

An interesting case is provided by an instance in which Macedonians identify themselves as Greeks and speakers of Greek. The Roman historian Livy (first centuries BC and AD), writing of events in the war waged by Philip the Fifth of Macedon and his Arcarnanian Greek allies against Athens, with Rome as its own ally, records a meeting of the council of the Aetolian Confederacy, at which representatives from Philip, from Athens and from Rome address the council, each seeking Aetolian assistance in the war (200 BC). In his speech to the council,

the Macedonian ambassador refers to the Romans as “a foreign people set apart more by language and customs and laws than by the space of sea and land” (31.29.12). In contrast, “Aetolians, Acarnanians and Macedonians [are] people of the *same language* . . . [and] with foreigners, with barbarians *all Greeks* are, and will be, at eternal war” (31.29.15). The dialect of the Aetolian Confederacy, a league of the Aetolians of northwest Greece, was the Northwest Greek Koine, a “common” dialect used throughout regions controlled by the Confederacy (see Ch. 25, §1.1.5). Is it this lingua franca to which Livy has his Macedonian diplomat self-servingly refer? One could well imagine that it would be the Macedonian’s *langue de choix* on such an occasion. The Acarnanians also inhabited northwest Greece, though Acarnanian inscriptions from this period are written in the Doric Koine, only slightly different from the Aetolian dialect.

Surviving Macedonian texts have not proved helpful in identifying the native language of the Macedonians. Most of the Macedonian inscriptions are written in Attic Greek, the dialect broadly disseminated by Philip and Alexander. A fourth-century BC inscription found recently in the remains of the great Macedonian city of Pella appears to be written in a variety of Northwest Greek and has led to conjectures that this may be the previously unattested Macedonian language (see the comments of Brixhe and Panayotou 1994b:209 along with the mention of other finds in n.19).

The evidence provided by Macedonian glosses is conveniently summarized by Katičić (1976:108–112), who analyzes these as belonging to three different classes. One class consists of words that are quite close to known Greek lexemes, some, though probably not all, of which appear likely to be loanwords directly from Greek: for example, *kommárai*; compare Greek *kámmaroi* (κάμμοροι), a type of lobster (pl.). A second set is made up of Macedonian words which have no Greek counterparts, such as *aliē* “boar.” The third group is similar to the first to the extent that it consists of Macedonian words apparently having Greek counterparts; it differs from the first class, however, in that these Macedonian words are perhaps to be analyzed as cognates of the Greek lexemes, rather than borrowings. In other words, by such an analysis, the related Macedonian and Greek forms have evolved historically from words occurring in a common parent language, either Proto-Indo-European or, alternatively, some later, intermediate Balkan Indo-European language. Compare, for example, Macedonian *adē* “sky” and Greek *aitḗr* (αἰθήρ); Macedonian *kebalá* “head” (cf. *gabalá* which the Greek lexicographer Hesychius also glosses as “head,” without identifying the linguistic source of the word) and Greek *kepḥalē* (κεφαλή). If such sets are rightly analyzed as cognates, the Macedonian language departs conspicuously from Greek in showing voiced unaspirated rather than voiceless aspirated reflexes of the earlier Indo-European voiced aspirated stops (on the Greek development, see Ch. 24, §3.7.1).

1.3.7 Messapic

The Messapii were a people of southeast Italy, inhabiting ancient Calabria (the Sallentine peninsula, the “heel” of the Italian “boot”). Strabo, the Greek geographer, records (*Geography* 6.3.1) that the Greeks give the name *Messapia* to that region, also called *Iapygia*, but adds that the locals of the area make a distinction between the Salentini (in the south) and the Calabri. Northward lies the country of the Peucetii and of the Daunuii (Apulia). For Polybius (*Histories* 3.88.4), however, Iapygia is the region inhabited by the Daunuii, Peucetii, and Messapii (though elsewhere he writes of “Iapyges and Messapii”; see *Histories* 2.24.11).

Messapic survives in a large number of inscriptions, recording chiefly proper names, dating from about the sixth to the first century BC (the most abundantly attested ancient language not to receive individual treatment in this encyclopedia), including many recent

finds from a grotto in Lecce (see Santoro 1983–1984). This language of ancient Italy is Indo-European, but not Italic; that is, it is not a member of the subfamily to which belong Latin and Sabellian (see Chs. 32 and 33). No close genetic affiliation with any other known Indo-European language can be definitively demonstrated, though a close connection to Illyrian has been alleged. Indeed, the Messapic materials provided a major component of the evidence adduced by Krahe and others for the study of Illyrian. There do exist ancient traditions about the settling of southeast Italy by Illyrian peoples. For example, Pliny (*Natural History* 3.102) makes cursory reference to the story that the “Paedicali” of Apulia were descended from nine young men and nine young women of Illyria. A linking of the two languages, Illyrian and Messapic, must, however, remain a linguistically unverifiable hypothesis until such time as Illyrian is better attested.

1.3.8 Zapotec

Far away from Italy and the Balkans, in Mesoamerica, yet another language of antiquity is attested. Zapotec is one of several documented early Mesoamerican languages, others being Mayan (Ch. 43), Epi-Olmec (Ch. 44), Mixtec and Aztec; both of the last-named are attested by about AD 1100 but are best known from sixteenth-century AD manuscripts (*inter alia*, for Mixtec, see Marcus 1992:57–67 and Jansen 1992:20–33; for Aztec, see Marcus 1992 and Prem 1992:53–69; on the pictographic records of the Tlapanecs, see Vega Sosa 1992:34–52). Zapotec inscriptions, carved in stone like those of the Mayans and Epi-Olmecs, may date as early as 500 or 600 BC (though the earliest uncontroversial dates are between 400 and 200 BC) and are last attested in about AD 900 (as with Mixtec and Aztec, Zapotec manuscripts also occur in the sixteenth century, though the corpus is small). Several dozen short inscriptions exist, as well as a large number of calendrical citations, providing perhaps one to three hundred distinct glyphic components.

Owing to the difficulty in obtaining information on this language, a brief grammatical sketch of Zapotec, based on our present, limited understanding of the language, has been included as an appendix to Chapter 44.

1.4 Format and conventions

Each chapter, with only the occasional exception, adheres to a common format. The chapter begins with an overview of the history (including prehistory) of the language, at least up to the latest stage of the language treated in the chapter, and of those peoples who spoke the language (§1, HISTORICAL AND CULTURAL CONTEXTS). Then follows a discussion of the development and use of the script(s) in which the language is recorded (§2, WRITING SYSTEMS); note that the complex Mesopotamian cuneiform script, which is utilized for several languages of the ancient Near East – Sumerian (Ch. 2), Elamite (Ch. 3), Hurrian (Ch. 4), Urartian (Ch. 5), Akkadian and Eblaite (Ch. 8), Hittite (Ch. 18), Luvian (Ch. 19) – and which provides the inspiration and graphic raw materials for others – Ugaritic (Ch. 9) and Old Persian (Ch. 28) – is treated in most detail in Chapter 8, §2. The next section presents a discussion of phonological elements of the language (§3, PHONOLOGY) identifying consonant and vowel phonemes, and treating matters such as allophonic and morphophonemic variation, syllable structure and phonotaxis, segmental length, accent (pitch and stress), and synchronic and diachronic phonological processes. Following next is discussion of morphological phenomena (§4, MORPHOLOGY), focusing on topics such as word structure, nominal and pronominal categories and systems, the categories and systems of finite verbs and other verbal elements (for explanation of the system of classifying Semitic verb

stems – G stem, etc. – see Ch. 6, §3.3.5.2), compounds, diachronic morphology, and the system of numerals. Treatment of syntactic matters then follows (§5, SYNTAX), presenting discussion of word order and coordinate and subordinate clause structure, and phenomena such as agreement, cliticism and various other syntactic processes, both synchronic and diachronic. The description of the grammar closes with a consideration of the lexical component (§6, LEXICON); and the chapter comes to an end with a list of references cited in the chapter and of other pertinent works (BIBLIOGRAPHY).

To a great extent, the linguistic presentations in the ensuing chapters have remained faithful to the grammatical conventions of the various language disciplines. From discipline to discipline, the most obvious variation lies in the methods of transcribing sounds. Thus, for example, the symbols *ś*, *ṣ*, and *ṭ* in the traditional orthography of Indic language scholarship represent, respectively, a voiceless palatal (palato-alveolar) fricative, a voiceless retroflex fricative, and a voiceless retroflex stop. In Semitic studies, however, the same symbols are used to denote very different phonetic realities: *ś* represents a voiceless lateral fricative while *ṣ* and *ṭ* transcribe two of the so-called emphatic consonants; the latter a voiceless stop produced with a secondary articulation (velarization, pharyngealization, or glottalization), the former either a voiceless fricative or affricate, also with a secondary articulation. Such conventional symbols are employed herein, but for any given language, the reader can readily determine phonetic values of these symbols by consulting the discussion of consonant and vowel sounds in the relevant phonology section.

Broad phonetic transcription is accomplished by means of a slightly modified form of the International Phonetic Alphabet (IPA). Most notably, the IPA symbols for the palato-alveolar fricatives and affricates, voiceless [ʃ] and [tʃ] and voiced [ʒ] and [dʒ], have been replaced by the more familiar [š], [č], [ž], and [j] respectively. Similarly, [ɣ] is used for the palatal glide rather than [j]. Long vowels are marked either by a macron or a colon.

In the phonology sections, phonemic transcription, in keeping with standard phonological practice, is placed within slashes (e.g., /p/) and phonetic transcription within square brackets (e.g., [p]); note that square brackets are also used to fill out the meaning of a gloss and are employed as an element of the transcription and transliteration conventions for certain languages, such as Elamite [Ch. 3] and Pahlavi [Ch. 30]). The general treatment adopted in phonological discussions has been to present transcriptions as phonetic rather than phonemic, except in those instances in which explicit reference is made to the phonemic level. Outside of the phonological sections, transcriptions are usually presented using the conventional orthography of the pertinent language discipline. When potential for confusion would seem to exist, transcriptions are enclosed within angled brackets (e.g., <p>) to make clear to the reader that what is being specified is the *spelling* of a word and not its *pronunciation*.

Bibliography

- Bader, F. (ed.). 1994. *Langues indo-européennes*. Paris: CNRS Editions.
- Bricker, V. 1992. *Epigraphy*. Supplement to the Handbook of Middle American Indians, vol. V. Austin: University of Texas Press.
- Brixhe, C. and A. Panayotou. 1994a. "Le thrace." In Bader 1994, pp. 179–203.
- . 1994b. "Le macédonien." In Bader 1994, pp. 205–220.
- Brown, E. 1990. "Traces of Luwian dialect in Cretan text and toponym." *Studi micenei ed egeo-anatolici*, pp. 225–237.
- Chadwick, J. 1990. "Linear B and related scripts." In *Reading the Past*, pp. 137–195. Introduction by J. Hooker. Los Angeles/Berkeley: University of California Press; London: British Museum.

- Conway, R., J. Whatmough, and S. Johnson. 1933. *The Prae-Italic Dialects of Italy* (3 vols.). Cambridge, MA: Harvard University Press.
- Cowgill, W. and M. Mayrhofer. 1986. *Indogermanische Grammatik*, vol. I. Heidelberg: Carl Winter.
- Daniels, P. 1996. "The first civilizations." In Daniels and Bright 1996, pp. 21–32.
- Daniels, P. and W. Bright (eds.). 1996. *The World's Writing Systems*. Oxford: Oxford University Press.
- Dhorme, E. 1946. "Déchiffrement des inscriptions pseudo-hiéroglyphiques de Byblos." *Syria* 25:1–35.
- Diringer, D. 1968. *The Alphabet: A Key to the History of Mankind*, vol. I. New York: Funk and Wagnalls.
- Duhoux, E. 1982. *L'Étéocrétois*. Amsterdam: Gieben.
- Dunand, M. 1945. *Byblia grammata: Documents et recherches sur le développement de l'écriture en Phénicie*. Beirut: Ministry of Education of Lebanon.
- Englund, R. 1996. "The Proto-Elamite script." In Daniels and Bright 1996, pp. 160–164.
- Finkelberg, M. 1990–1991. "Minoan inscriptions on libation vessels." *Minos* 25–26:43–85.
- Griffith, F. 1911. *Karanog: The Meroitic Inscriptions of Shablul and Karanog*. Philadelphia: University Museum.
- . 1912. *Meroitic Inscriptions II*. London: Egypt Exploration Fund.
- Jackson, K. 1980. "The Pictish language." In F. Wainwright (ed.), *The Problem of the Picts*, pp. 129–160. Perth: Melven.
- Jansen, M. 1992. "Mixtec pictography: conventions and contents." In Bricker 1992, pp. 20–33.
- Katičić, R. 1976. *Ancient Languages of the Balkans*, part 1. The Hague: Mouton.
- Krahe, H. 1940. "Der Anteil der Illyrier an der Indogermanisierung Europas." *Die Welt als Geschichte* 6:54–73.
- . 1955. *Die Sprache der Illyrier* (2 vols.). Wiesbaden: Otto Harrassowitz.
- Lehmann, W. 1967. *A Reader in Nineteenth-Century Historical Indo-European Linguistics*. Bloomington: Indiana University Press.
- Mallory, J. 1989. *In Search of the Indo-Europeans*. London: Thames and Hudson.
- Marcus, J. 1992. *Mesoamerican Writing Systems*. Princeton: Princeton University Press.
- Masson, E. 1974. *Cyprominoica*. Göteborg: Paul Åströms Forlag.
- . 1977. "Présence éventuelle de la langue hourrite sur les tablettes chyro-minoennes d'Enkomi." *Revue Roumaine de Linguistique* 22:483–488.
- Masson, O. 1983. *Les inscriptions chypriotes syllabiques*. Paris: Édition E. de Boccard.
- McMannus, D. 1991. *A Guide to Ogam*. Maynooth: An Sagart.
- Palaima, T. 1988. "The development of the Mycenaean writing system." In J. Olivier and T. Palaima (eds.), *Texts, Tablets and Scribes*, pp. 269–342. Supplement to *Minos* 10.
- . 1989. "Cypro-Minoan scripts: Problems of historical context." In Y. Duhoux, T. G. Palaima, and J. Bennet (eds.), *Problems in Decipherment*, pp. 121–187. Louvain-la-Neuve: Peeters.
- Palmer, L. 1968. "Linear A and the Anatolian languages." In *Atti e memorie del 1° congresso internazionale di micenologia*, vol. I, pp. 339–354. Rome: Edizioni dell'Ateneo.
- Parpola, A. 1994. *Deciphering the Indus Script*. Cambridge: Cambridge University Press.
- . 1996. "The Indus Script." In Daniels and Bright 1996, pp. 165–171.
- Poultney, J. 1979. "The language of the Northern Picene inscriptions." *Journal of Indo-European Studies* 7:49–64.
- Prem, H. 1992. "Aztec writing." In Bricker 1992, pp. 53–69.
- Pulgram, E. 1978. *Italic, Latin, Italian*. Heidelberg: Carl Winter.
- Santoro, C. 1983–1984. *Nuovi studi messapici*. Galatina, Italy: Congedo.
- Swiggers, P. 1996. "The Iberian Scripts." In Daniels and Bright 1996, pp. 108–112.
- Untermann, J. 1975. *Monumenta Linguarum Hispanicarum i, Die Münzlegenden*. Wiesbaden: Dr. Ludwig Reichert.
- . 1980. *Monumenta Linguarum Hispanicarum ii, Die Inschriften iberischer Schrift aus Südfrankreich*. Wiesbaden: Dr. Ludwig Reichert.
- . 1990. *Monumenta Linguarum Hispanicarum iii, Die iberischen Inschriften aus Spanien*. Wiesbaden: Dr. Ludwig Reichert.
- . 1997. *Monumenta Linguarum Hispanicarum iv, Die tartessischen, keltiberischen und lusitanischen Inschriften*. Wiesbaden: Dr. Ludwig Reichert.

- Vega Sosa, C. 1992. "The annals of the Tlapanecs." In Bricker 1992, pp. 34–52.
- Wenig, S. 1982. "Meroe, Schrift und Sprache." In W. Helck and W. Westendorf (eds.), *Lexikon der Ägyptologie*, vol. 4, pp. 104–107. Wiesbaden: Otto Harrassowitz.
- Whittaker, G. 1992. "The Zapotec writing system." In Bricker 1992, pp. 5–19.
- Wilkes, J. 1992. *The Illyrians*. Oxford: Blackwell.
- Woodard, R. 1997. "Linguistic connections between Greeks and non-Greeks." In J. Coleman and C. Walz (eds.), *Greeks and Barbarians*, pp. 29–60. Bethesda, MD: CDL Press.

Hurrian

GERNOT WILHELM

1. HISTORICAL AND CULTURAL CONTEXTS

1.1 History of the language and its speakers

Hurrian is an ancient Near Eastern language widely spoken in the northern parts of the Fertile Crescent (present-day northern Iraq, northern Syria, southeast Turkey) from at least the last quarter of the third millennium BC on until the end of the second millennium BC. It survived for another half millennium in small pockets in the mountainous areas north of ancient Assyria.

A cognate language of Hurrian is Urartian (see Ch. 5) which is attested in texts from the late ninth to the late seventh century BC. Apart from Urartian, Hurrian is an isolated language without a genetic relation to any other known ancient Near Eastern language. A genetic relation between (reconstructed) Proto-Urarto-Hurrian and (reconstructed) Northeast Caucasian has been argued for, but it is not generally accepted. If the connection could be demonstrated, it would be a rather distant one.

Hurrian is first attested in a few words and personal or place names mentioned in Akkadian texts of the Akkade period (twenty-third to twenty-second centuries BC). The term *Old Hurrian* (herein abbreviated OH) has been coined for the language of a royal inscription most likely to be dated to the Ur III period (twenty-first to twentieth centuries BC), but it is also used for the more archaic dialect(s) of the second millennium.

During the first half of the second millennium BC (Middle Bronze Age) there are many hundreds of Hurrian personal names attested from the northern parts of the Fertile Crescent (from the Zagros Mountains in the east to the Mediterranean coast), but only little more than a dozen Hurrian texts, still unintelligible for the most part.

By far the majority of Hurrian texts comes from the second half of the second millennium BC (Late Bronze Age). Hurrian disappeared as a result of political and ethnic shifts occurring from the late fourteenth century BC onwards. Except perhaps in remote mountainous areas east of the upper Tigris, Hurrian became extinct during the Dark Ages, beginning in the twelfth century BC.

The modern name of the language (English “Hurrian,” French “hourrite,” German “hurritisch”) is based on the geographical term *Hurri* which is not very well defined (presumably denoting most of Upper Mesopotamia). On the basis of this name, Hittite forms an adjective *hurili* “Hurrian” (adjective in *-li* formed from *hur-la* “inhabitant of the land of Hurri”) which qualifies Hurrian language incantations used in Hittite rituals. In the so-called “Mittani letter” (fourteenth century BC, see §1.2), the Hurrian adjective *hurroge* (variant

hurvoḡe) "Hurrian" refers to the country; it is unknown whether it could also designate the language.

Other terms for the language are obsolete – *Mitanni* (based on the name of a country in Upper Mesopotamia); *Subarian* (based on the geographical term *Subir*, *Subartu*). The earliest Hurrian attestations and the linguistic relationship with Urartian point to an origin in the most northeastern parts of the Fertile Crescent and in the mountainous areas beyond (most northeastern Syria, most northern and northeastern Iraq, southeastern Turkey). A connection with the flourishing Transcaucasian Early Bronze culture is possible, but cannot be demonstrated.

The earliest city-states with Hurrian rulers, and presumably a population which at least in part spoke Hurrian, were under strong southern (Akkadian, Sumerian) cultural influence and military pressure. Already about 1800 BC there was a solid Hurrian element in the populations that lived between the Mediterranean and the Euphrates, most likely as a result of movements at the end of the third millennium BC.

At the end of the sixteenth century BC, the kings of Mittani (conventionally also "Mitanni") in Upper Mesopotamia united most of the Hurrian-speaking countries under their control. The dynasty preserved some archaic Indo-Aryan traditions of unknown origins (dynastic names, some gods known from Vedic sources, hippological terms). During the fifteenth century Mittani struggled with Egypt for the control of Syria west of the Euphrates. A balance was reached shortly after 1400 by a peace treaty and the beginning of a series of dynastic marriages. It is in this context that King Tušratta addressed the so-called Mittani Letter to Pharaoh Amenophis III. In the middle of the fourteenth century BC, Mittanian power declined rapidly as a consequence of dynastic turmoil and the rise of Hittite and Assyrian power; the Hittites conquered Mittani's vassal states west of the Euphrates, whereas the east was annexed to Assyria, though the Mittanian dynasty was able to keep control of a part of its former empire for four more generations. Eventually, however, it disappeared in the course of Assyrian military expansion. The Assyrians removed whole population groups from former Mittani and settled Assyrians in their place in order to gain better control of the region. This policy undoubtedly accelerated the disappearance of the Hurrian language.

By about 1400 BC, the Hittite dynasty had already adopted cultural traditions from the Hurrian-speaking parts of southern Anatolia (Kizzuwatna). Consequently, Hittite kings supported Hurrian cults and introduced them into their capital of Hattuša and in several north Anatolian provincial centers (Sapinuwa, Samukha). Between 1400 and the Hittite collapse, Hurrian thus became a language of cult and learning far removed from Hurrian-speaking areas.

1.2 Sources

The oldest Hurrian text is the royal inscription of Tišatal, *endan* of Urkeš. The texts from c. 2000–1500 BC are mainly short incantations from places outside Hurrian-speaking areas – Babylonia (Larsam?) and the Middle Euphrates (Mari, Tuttul) – but there also exist a few texts of other, not yet identified genres (from Mari, Tigunanu).

The most important source for the study of Hurrian up to the present time has been the Mittani Letter written about 1355 BC. This diplomatic document of nearly five hundred lines was discovered in the Egyptian capital of Amarna in 1887. All the other known messages sent to Egypt by the royal court of Mittani are written in Akkadian, which at that time served as the diplomatic vernacular throughout the ancient Near East. Archeology has not yet been very successful in uncovering cuneiform tablets in Mittani proper; though recent finds from Tell Brak on the Khabur river show that Akkadian was widely used in the area, especially for

deeds (there is also a small fragment of a Hurrian letter). The thousands of texts found at Nuzi (close to Kirkuk, northern Iraq) and Alalakh (close to Antioch on the Orontes) are all written in Akkadian, but often display Hurrian influence on the levels of lexicon, grammar, and personal names (anthroponymy).

The trade center Ugarit on the Mediterranean coast has yielded a small but important and diverse number of Hurrian texts. A bilingual Sumero-Hurrian lexical list displays unusual Hurrian forms. Several other tri- or quadrilingual (Sumerian, Akkadian, Hurrian, and, optionally, Ugaritic) lists of words or divine names (theonyms) as well as a short bilingual (Akkadian, Hurrian) wisdom text have made important contributions to our understanding of the Hurrian lexicon. Some Hurrian texts are written in the Ugaritic alphabetic script. A group of Hurrian cult songs combine a (mostly unintelligible) text with musical terms (based on Akkadian ones). One single Hurrian letter suggests the use of the language also outside the sphere of cult and learning.

At Emar on the middle Euphrates omen texts and a trilingual (Sumerian, Akkadian, Hurrian) god list were found in the 1970s, but by 1999 they remained unpublished.

By far the majority of Hurrian texts come from Hittite libraries. The most important belong to a series of bilingual (Hurrian, Hittite) literary texts, including a myth, a historico-religious tale, and wisdom texts. Among other texts, Hittite rituals with Hurrian incantations or offering-lists stand out, but there are also epics, myths, prayers, and omen texts. Many Hurrian texts are reported to have been found at Sapinuwa (Ortaköy south of Çorum) beginning in 1991; they are also still unpublished.

1.3 Dialects

Despite its vast geographical distribution and its attested history of about a millennium, Hurrian is remarkably homogeneous. The two main dialects are that of the Mittani Letter and the dialect (or presumably a group of closely related dialects) called *Old Hurrian* (being much earlier attested than Mittani Hurrian). Old Hurrian is closer to Urartian, which seems to have separated from Proto-Urarto-Hurrian not later than the early second millennium BC. It is also the dialect on which the study of Hurrian proper names rests. The main features of Old Hurrian have become clearer only since 1983 when the above-mentioned Hurro-Hittite bilingual was discovered. The chief differences between the two dialects lie in the verbal system and in syntax, though the much more complicated syntax of the Mittani letter (virtually our only source for Mittani Hurrian) and its wider use of enclitics may be due to the demands of diplomacy. A few dialectal differences within Old Hurrian are discernible.

2. WRITING SYSTEMS

2.1 Syllabic cuneiform

Hurrian was mainly written in the syllabic cuneiform script of Akkadian. Departing from common Akkadian spelling practice (see Ch. 8, §2), only a few logograms (word signs originally used to write Sumerian, hence also called *Sumerograms*) were used in writing Hurrian texts.

The scribe of the Mittani Letter used a restricted inventory of syllabic symbols (41 CV signs, 31 VC signs, and 26 CVC signs, some of which had two values – *ḫar/ḫur*, *kal/tan*). The Akkadian script adopted for writing Hurrian distinguished only partially between /e/ and /i/;

the Mittani Letter makes full use of the given oppositions (*te/ti*, *še/ši*, *me/mi*, *en/in*, *el/il*). Long vowels are rendered by *plene-spelling* (e.g., *še-e-*, *ta-a-*). The two vowel signs *u* and *ú* are strictly distinguished in the Mittani Letter and in some texts from Hattuša, indicating a phonemic distinction of /o/ versus /u/.

The syllabary of the Mittani Letter does not distinguish (at least in some cases) between voiced and voiceless stops, but utilizes one Akkadian sign out of a pair – *pa*, not *ba*; *ta*, not *da*; *du*, not *tu*; and so forth. In two instances, the script of the Mittani Letter redefines a pair of Akkadian signs: *gi* and *ki* encode a difference no longer in consonantal voicing, but in vowel quality – *gi* is used for /Ke/ and *ki* for /Ki/. Correspondingly, *gu* represents /Ku/ and *ku* /Ko/.

The sign *wa* can be used for a labiodental fricative plus any vowel; in texts from Hattuša a small vowel sign is added in order to facilitate the correct reading.

Going a step beyond Akkadian practice, Hurrian scribes repeated a vowel sign in word-initial position before a single consonant in order to represent a long vowel: for example, *u-u-mi-i-ni* for *ōmīni*.

2.2 Alphabetic cuneiform

Some of the texts from Ugarit are written in the Ugaritic cuneiform consonantal (so-called alphabetic) script, presumably by Ugaritic-speaking scribes. These yield important evidence for the phonology and phonetics of Hurrian, as the consonantal script encodes differences in consonants which are obscured by the syllabic script.

3. PHONOLOGY

3.1 Consonants

Since Hurrian was written with scripts which were designed for other languages, it is difficult – to a degree even impossible – to establish the phonemic inventory of Hurrian. In writing Hurrian words and names, non-Hurrian scribes in Babylonia and Ugarit distinguish between voiced and voiceless consonants in keeping with their own native phonologies. However, the distribution of voiced and voiceless consonants in Hurrian follows a strictly positional pattern – in other words, is allophonic. Obstruents are always perceived as voiceless (i) in word-initial position; (ii) in intervocalic position when long (doubled); and (iii) in contact with another consonant except the sonorants /m/, /n/, /l/, and /r/. Conversely, obstruents are voiced in all other positions: (i) when word-final; (ii) in intervocalic position when short (single); and (iii) in contact with /m/, /n/, /l/, and /r/. The resulting consonantal inventory would then appear to be as follows, with capital letters used noncommittally (“archiphonemically”) to transcribe the obstruents displaying allophonic voicing:

(1) Hurrian consonantal phonemes

P	T	K	
		ʈʂ	
F	Š	S	Ḫ
m	n		
	l	r	
w	y		

The affricate /tʃ/ (transcribed as *c*) is uncertain. Though voicing is not phonemic, it is by convention (following E. A. Speiser and I. M. Diakonoff) marked in (bound) transcriptions in order to facilitate research on loanwords into and from Hurrian (*p:b, t:d, k:g, f:v, s:z, š:ž, ḥ:ğ*). Note that also according to convention, the so-called broad transcription (transliteration) of syllabic cuneiform uses single bars (-) to separate syllabic symbols, whereas the morphemic transcription uses double bars (=).

3.1.1 Obstruents

Since the inventory of (1) is certainly too small, we have to assume that there were two or more sets of obstruents with different phonemic manners of articulation which remain unknown.

The fricative /F/ appears to be a labiodental, as the Mittani Letter distinguishes /F/ and bilabial /w/, the first one written with the sign *wa*, the second one with *ú*. The phonetic realization of /S/ and /š/ is unknown; the latter is rendered as an interdental fricative – [θ] or [ð], depending on position – by Ugaritic scribes.

The texts from Hattuša often replace /P/ by a fricative, apparently in all positions. Whether this fricative is identical with the /F/ of the Mittani Letter or different (bilabial) is unknown.

3.1.2 Sonorants

The bilabial glide /w/ appears in word-internal and final position; in word-initial position it seems to be restricted to loanwords and foreign names. There is also a glide /y/ which, however, is rendered as *i*, *ī* in modern transcription, as the writing system in many cases is ambiguous.

3.2 Vowels

The Mittani Letter distinguishes five phonemic vowels – /a/, /e/, /i/, /o/, /u/ – with two quantities each. Only the Mittani Letter carefully marks vowel length. The texts from Hattuša show an instability of the opposition /e/ : /i/.

The existence of diphthongs is uncertain. At least some sequences of vowels (e.g., the suffix *-ae*, see §4.4.9, instrumental) can be shown as bisyllabic, and may be even divided by a glottal stop.

3.3 Phonotaxis

The practice of syllabic cuneiform orthography prohibits the unambiguous representation of biconsonantal clusters in word-initial or final position, and of triconsonantal clusters in word-internal position. There are no hints that such clusters actually exist in Hurrian; moreover, the appearance of anaptyctic vowels suggests that in this respect the script conforms to the language.

The liquids /l/ and /r/ do not appear in word-initial position.

In the language of the Mittani Letter, strict constraints govern final position: vowels or /n/ occur in most cases; the consonants /š/, /w/, and /F/ or /P/ are limited to one suffix each. In Old Hurrian – especially in the case of divine and place names – /T/, /K/, /Ḫ/, /l/, /m/, and /r/ also appear in final position.

3.4 Accent

Hurrian seems to have a stress accent which falls on the penultimate syllable of words (including their suffixes), enclitics not counted. In some cases, stress causes a vocalic change (lengthening and lowering): for example, *túri* "low" versus *turé=na* (this is the morphemic transcription; the transliteration of the cuneiform spelling is *du-ú-re-e-na*) "the low ones."

3.5 Phonological processes

Several Hurrian phonological processes, synchronic and/or diachronic, can be identified.

3.5.1 Anaptyxis

Vowels are inserted under two conditions:

1. Presumably with the shift of stress caused by addition of a suffix: for example, (i) *évri* "lord": *evérni* "king"; (ii) *talógli* "servant": *talógól=la* (pl.); (iii) *havúrni* "heaven": *havurún=nē=ž* (erg.); (iv) *am=om=i=nni* "administrator": *am=om=i=nin=n(a)=až=i?=na* (gen. pl., double stress?).
2. With the *-n* affix of the jussive (see §4.5.12.2) and ablative (see §4.4.9) before enclitic personal pronouns, except that of the third-person singular: (i) *haž=i=en* "may he listen": *haž=i=en=i=ll(a)=ān* "may he listen to them"; (ii) *ed(i)=ī=dan* "because of": *ed(i)=ī=dan=i=lla=man* "by himself... them" (but *ed(i)=ī=da(n)=nna=man* "he... by himself").

3.5.2 Segment loss

The regular disappearance of sounds is seen in three contexts:

1. In the morphologically conditioned contact of two vowels, the first one is elided: for example, (i) *šēna* "brother": *šēniffu-* "my brother"; (ii) *fīradi* "nobleman": *fīradardi* "nobility" (in morphemic transcription the elided vowel is given in brackets: *šēn(a)=iffu-*, *fīr=ad(i)=ardi*). For an exception see §3.5.3.
2. The vowels /a/ and /i/ are syncopated between (simple) /n/, /r/, /l/ and (archaic?) dental stops: **kul=i=l=e* → *kulle* "I should like to say"; **ēni=na* → *ēnna* "the gods"; **kud=id=e(n)* → *kut=t=e(n)* "may they fell."
3. The consonantal segment of the genitive suffix *-ve* and the dative *-va* is lost after the plural suffix *-až* (see §4.4.9).

3.5.3 Vowel contraction

The contact of the final (short) /a/ of the enclitic pronouns (see §4.4.10.1) and the initial (short) /a/ of the enclitic connective *-an* results in a long vowel: *-tt(a)=ān* spelled *-Vt-ta-a-an*.

3.5.4 Assimilation

Hurrian shows both consonant and vowel assimilation, progressive and regressive:

1. When in the case mentioned in §3.5.2, 2, the two consonants are different, the second one is assimilated to the first one: **avari=ne-* → *avarre* "field"; **tād=ugār=i=l=eva* → *tadugarreva* (see §4.5.12.6).

2. The consonantal segment of the genitive suffix *-ve* and dative *-va* is assimilated to a preceding /P/, /T/, or /Š/: *Tēššob* (a god), gen. *Tēššop=pe*; *Hebat* (a goddess), gen. *Hebat=te* (see, however, §3.5.5). It is partially assimilated to preceding /u/: *šēn(a)=iffu=we* “of my brother.”
3. Personal names composed of a verbal form and a divine name display various assimilations at the junction: for example, *Ag=i=p-Tēššob* → *Ag=i=t-Tēššob*; *Hud=i=p-Šimīga* → *Hud=i=š-Šimīga*.
4. The vowel of the two suffixes *-Všt-* (see §4.5.2, 2) and *-kkV* (see §4.5.7) assimilate to the preceding vowel (“vowel harmony”).

3.5.5 Metathesis

Consonantal stems joined with a suffix exhibit metathesis: for example, **kik=ši* → *kiški* “third”; *Kužaḫ=fe* → *Kužap/fḫe* “of the Moongod”; *Hebat=fi* → *Hebap/fti* “of (the goddess) Hebat.”

The dialect of the texts from Nuzi often (in some cases regularly) inverts the sequence *consonant + liquid*, especially when the initial consonant is a fricative: for example, *fağri*: *fağri*; *eğli*: *ełgi*; *evri*: *ervri*; *šadna*: *šanda*.

4. MORPHOLOGY

4.1 Word classes

Hurrian grammars distinguish the following word classes: nouns, adjectives (mostly derived from nouns), pronouns, numbers, verbs, and particles (including enclitics). Nouns, numbers, and verbs may easily change their word class: for example, *eman* “ten,” *eman=am=ož=aw* “I made tenfold,” *eman=di* “group of ten,” *eman=d=o=g=li* “decurio”; *ḫan=i* “child,” *ḫan=ašt=i=kki* “she will not give birth,” *ḫan=ir(i)=ra* “those who have given birth,” *ḫan=o/u=mb=a=z=ḫe* “fertility”(?).

4.2 Roots

Hurrian words are composed of (i) roots, (ii) optional root-complements, and (iii) monofunctional nominal or verbal suffixes in a strictly sequential order. The root is always in initial position. Most roots are monosyllabic, but a few are reduplicated. The morphology of Hurrian is fundamentally of the agglutinating type.

4.3 Root-complements

A root can be semantically modified by one or two (possibly three) root-complements. In many cases, the semantic value of the root-complements has not yet been established. Root-complements in most cases are monosyllabic; those which traditionally have been defined as bisyllabic might well be composed of two root-complements. Root-complements are listed in (2), though the list is not exhaustive.

(2) Hurrian root-complements

- A. *-ağ-* (*-aḫḫ-*?), *-iğ-*, *-oğ-*, *-uğ-*, meaning unknown: *tapš=ağ=i* “cupbearer,” *pūz=iğ-* “dip into,” *irn=oğ-* “make equivalent,” *šab=uğ-* “?”

- B. -*al*-, meaning unknown: *samm=al*- "tear off," *hež=āl*- "be naked"
- C. -*am*-, factitive: *eman=ām*- "make tenfold," *šin=am*- "double," *nikkass=am*-, *nissakk=am*- "account" (Akkadian *nikkassu* "account")
- D. -*an*-, -*ann*-, causative: *keb=ān*- "send," *ar=ann*- "let give," *an=an*- "please"
- E. -*an*-, meaning unknown: *ḫab=an*- "go," *kil=ān*- "?"
- F. -*and*-, meaning unknown (bimorphemic?): *pic=and*- "rejoice"
- G. -*ang*-, meaning unknown (bimorphemic?): *pir=ang*- "flee"; *pūd=ang*- "report (to the authorities)"
- H. -*apš*-, -*epš*-, meaning unknown: *šin=apš*- "change" (*šin* "two"), *kig=apš*- "change repeatedly" (*kig(a)* "three"), *par=apš=i* a qualification of a field, *pur=apš=i* a priest, *tag=apš=i* "horse blanket," *eğ=epš*- "constrict"
- I. -*ar*-, iterative-frequentative: *am=ar*- "treat badly," *an=an=ar=ešk=i* "joy"(?). *ḫāž=ar*- "anoint," *ḫāž=ar=i* "oil," *pašš=ar*- "send (regularly)," *šid=ar*- "curse constantly," *šid=ar=ni* "curse" (noun), *tād=ar*- "love constantly," *fand=ar=i=nni* "cook," *urb=ar=i=nni* "butcher"
- J. -*až*-, intensive(?): *ḫaž=až*- "do listen" (?)
- K. -*om*-, meaning unknown: *am=om*- "supervise" (*am*- "look at something"), *am=om=i=nni* "chief administrator"
- L. -*ugar*-, reciprocal (bimorphemic?, cf. -*ar*-): *ag=ugar*- "dispatch," *ašt=ugar=i* "equivalent," *Ḫub=ušt=ugar=a* name of a divine vizier, *tād=ugār*- "love one another"
- M. -*up(p)*-, meaning unknown: *kad=upp*- "?" (*kad*- "say"), *tān=upp*- "?" (*tān*- "do, make")
- N. -*ur*-, meaning unknown: *ag=ūr=ni* "chiseling," *kul=ūr*- "spell over something" (*kul*- "say, speak")

4.4 Nominal morphology

The Hurrian noun (and a small number of nonderived adjectives) consists of (i) a root, which may be semantically modified by (ii) a root-complement or two, (iii) an optional noun-formation suffix, and in most cases (iv) a thematic vowel. In addition, by attachment of (v) derivational suffixes a noun may form a derived noun or adjective. There are two numbers (singular and plural), but no grammatical genders. To a noun (derived or primary) relational, possessive, number, case, and congruence suffixes – in a strictly sequential order – may be added, which may be followed by enclitics.

4.4.1 Nominalization of the root

A root may become a noun by addition of the thematic vowel -*i* (presumably -*ə* in final position, -*ē*- before a suffix, also -*e*?, see §4.4.3): thus, *eğl=i* "salvation" (*eğl*- "save"), *fur=i* "view; eye" (*fur*- "see"), *ḫan=i* "child" (*ḫan*- "give birth"), *ḫalv=i* "enclosure" (*ḫalv*- "enclose [by wall or fence]"), *ḫezm=i* "girdle" (*ḫezm*- "gird"), *ḫāž=ar=i* "oil" (*ḫāž=ar*- "to anoint repeatedly"), *mad=i* "wisdom" (*mad*- "to be, to prove oneself wise"), *sull=i* "fetter" (*sull*- "bind").

4.4.2 Noun-formation suffixes

Each of the following suffixes can be identified:

1. -(*a*)*d=i*, basic meaning unknown, in some cases collective: the allomorph -*di* appears when the root ends in a vowel or single post-vocalic *l*, *m*, or *n*, otherwise -*adi*:

amm=adi "grandfather, ancestor, elder," *šaḡadn=adi* "halfshekel" (*še/aḡt*- "half"), *šigl=ade* "shekel" (Akkadian *šiqlu*), *pariss=ade* a measure of capacity (Akkadian *parisu*), *kel=di* "luck, well-being," **ḡel=di* "sublime," *kum=di* "tower" (*kum*- "erect"?), *fir=adi* "nobleman" (**fir*- "remove, untie"), *nakk=adi* a form of real estate (*nakk*- "release"), *eman=di* "group of ten" (*eman* "ten"), *tumn=adi* "with four spokes" (*tumni* "four").

2. *-arb*, adjectives denoting age of animals (see §4.7.2, 6).

3. *-ardi*, collectives: *att=ardi* "forefathers" (*attai* "father"), *ēl=ardi* "female relatives" (*ēla* "sister"), *fir=ad=ardi* "nobility," *irīn=n(i)=ardi* "class of equals," *mariyā=nn(i)=ardi* "class of chariot fighters," *pura=m(e)=ardi* "domestic staff," *šāl=ardi* "group of daughters."

4. *-aure*, patient-oriented participle: *ḡuž=aure* "someone who is bound" (i.e., "a prisoner"; see also *-iri*).

5. *-bade*, meaning unknown: *ḡir=i=bade* "fixed by a peg" (*ḡiri* "wood"?), *tid=i=bade* "counting"; compare *-o/ubade*, negative adjectives (morphology unclear, with negative suffix *-ōv*- or derivational *-o*- [see 4.4.5]?): *nir=o/ubade* "bad" (*nir*- "be good"), *faḡr=o/ubade* "ugly, bad," *kul=o/ubade* "unnamed," *naḡḡ=o/ubade* "uninhabited."

6. *-danni*, *-denni* (*-da/e+nni* ?), OH *-dan*, terms of profession: *abul=dann*- "gate-keeper" (Akkadian *abullu* "city gate"), *ḡāz=ar=denn*- "perfume maker" (*ḡāz=ar=i* "oil"), *šellin=dann*- "administrator," *en=dan* title of a ruler (Sumerian *en*).

7. *-i(=)di*, meaning unknown: *šug=idi* "one horse carriage" (?) (from *šugi* "one"?), *tar=idi* "pot" (*tari* "fire"), *ḡub=idi* "young male calf" (from *ḡub*- "smash, break"?), presumably **pašš=idi* as base for *paššitḡe* "ambassador" (*pašš*- "send").

8. *-i(=)ri*, agent-oriented participle (cf. also *-aure*): *tab=iri* "someone who has cast (metal)," *pa=iri* "someone who has erected (a building)."

9. *-ki*, meaning unknown: *fut=ki* "son" (*p/fud*- "beget"), **katki* "utterance" (?) (*kad*- "speak"), *id=ar=gi* place for deposition of magically negative substances, *it=ki* "mortar" (*id*- "crush"), **ar=gi* "gift" (?) (*ar*- "give"). Presumably a group of nominals in *-a/e/i/oški* also contains this suffix *-ki*: *tād=ar=ašk(i)=ae* "affectionately" (?; a nominal used adverbially), *an=an=ar=eški* "joy" (?), *tād=ir=eški* "love" (?), *er=ōški* an object.

10. *-k(k)a*, meaning unknown: *Šav=oš=ka* name of goddess (older form *Ša(v)oža*, cf. *šav=ož=i* "great"), LUGAL-*ka*- "king," LÚ-*ka*- "man" (both based on Sumerograms), *aštaḡa* "woman" (*ašti* "woman"), *taḡḡag/kka* "man" (*taḡḡe* "man"); personal name *ḡanakka* (*ḡan=i* "child").

11. *-li*, nouns of profession (cf. also §4.4.6 (4A)): *kēb=li* "hunter" (*kēb*- "put," presumably traps), *tab=li* "smith" (*tav/b*- "to cast metal").

12. *-m(m)e*, meaning unknown: *pura=m(m)i/e* "slave"; perhaps also in *ulme/i* "female slave," *elami* "oath," *ḡalme* "singing," *ōlmi/e* "weapon," *tažme* "gift" (?).

13. *-ni*, individualizing, basic meaning unknown: *everni* "king" (*evre* "lord"), *irīn=ni* "equal" (from **irn=i* "?," this from *irn*- "be equivalent"), *šukka=ni* "single" (*šukki* "one"), *Māžriā=ni* "Egyptian," *Mitta=ni* (from *Maitta=ni*, toponym based on personal name *Maitta*). The suffix often appears as a form, that is parallel to *a*-stems (see §4.4.3.2): *p/fabni*, *p/faba* "mountain" (note also *faban=ni* "mountain range"), *muž=ni*, *muž=a* "good order," *tiž=ni*, *tiž=a* "heart." The suffix is attached to kinship terms and then used in the formation of personal names: *šen=ni* (*šēn=a* "brother"), *el=li* (*ēl=a* "sister"), *men=ni* (*men=a* female relative, twin sister[?]), *atta=ni* (*atta=i* "father," for *-i* see §4.4.4). Often *-ni* is suffixed to bi- or trisyllabic roots, which in some cases can be analyzed as *root plus root-complement*; the basic function remains unclear: *šid=ar=ni* "curse" (*šid*- "curse," with iterative *-ar*-), *kapp=ar=ni* a vessel (*kapp*- "fill"), *ḡavur=ni* "heaven," *šugur=ni*, *šēgur=ni* "life," *taržuwa=ni* "man." Compare also *-ni/-nni* as a derivational suffix (§4.4.6, 2).

14. *-šari* (*-zari* after *n*), collectives: *en=zari* "gods" (*eni* "god"), *furul=z/šari* "temple complex" (?) (*furli* "temples"), *hanizari* "children" (?) (*hani* "child"), *mariya=n=zari* "corps of chariot owners" (?) (*mariyanni* "chariot owner"), *tip=šari* "matter" (*tivi* "word, matter").

15. *-(a)=šše, -ži, -zi*, abstracts, but in some cases concrete nouns, especially words for buildings; also used for forming ordinal numbers (see §4.7.1). The allomorphs *-ži* and *-zi* appear after single postvocalic *m, n, l*, and *r*. The abstracts in *-a=šše* are exclusively derived from words for high-ranking men or for gods, which often are stems ending in *-a* (see §4.4.3.2).

15A. *Abstract nouns*: *all=a=šše* "queenship" (*alla=i* "queen"); **att=a=šše* "position of a father" (*atta=i* "father"), *puram=ži* "slavery" (*purame* "slave"), *šarr=a=šše* "kinship" (*šarri* "king"), *taḫḫ=a=šše* "manliness" (*taḫḫe* "male," < **taḫḫai*??), *taržuwan=zi* "mankind," *ušt=a=šše* "heroism" (*ušta=i* "hero"), also *ušt=a=n(i)=zi* (*uštani* "hero"), *tamga/ir=a=šše* "gain" (from Akkadian *tamkāru* "merchant"), *itk=a=l=zi* "purity" (*itki* "pure, clean"); *tal=aḡ=o=l=zi* "attraction" (in a ritual of evocation).

15B. *Concrete nouns*: *salam=ži* "statue" (from Akkadian *šalmu* "statue"), *pidar=ži* "stable for cattle" (*pidari* "bull"), *tibiš=ši* "strawstack" (*tibni*, from Akkadian *tibnu* "straw"), *lippur=ži* a building (from **nippuri*?).

16. *-umme/i*, infinitives: *itt=ummi* "go," *faḡr=umme* "be in good relation," *udr=ummi* "protect"; directive in *-e*: *kur=ušt=umme=n(e)=e* "in order to dig."

4.4.3 Thematic stems

Hurrian nouns are classified as *thematic* or *athematic* according to the presence or absence of a thematic (stem) vowel. While stems in ancient Indo-European languages are similarly distinguished (see Ch. 17, §3.4), the Hurrian and Indo-European processes are quite distinct and should not be confused.

4.4.3.1 i-stems

Most nouns have the thematic vowel *-i*. Apparently it has no specific function except to nominalize the root. In many cases it can be shown that *-i* before a suffix is lengthened and lowered to *-ē-* (see §3.4). It is not clear to what extent there exist *e*-stems distinct from the (i) *i*-stems (see e.g. *aš-ḫē* "animal skin," *šiye* "water," *ku-un-kal-le-e* "broad-tailed sheep," and the personal name *Še-iš-we-e* [šēšfe "kid"]), and (ii) the word formation or derivational suffixes *-me*, *-šše*, and *-ḡe/-ḫḫe*.

4.4.3.2 a-stems

The thematic vowel *-a* marks kinship terms, some divine names, and a few other words. For most *a*-stems there is a form in which *-ni* replaces *-a* (see §4.4.2, 14). Examples are *šēna* "brother," *ēla* "sister," *šāla* "daughter," *nēra* "mother," *mēna* "twin sister(?)" (see also §4.4.5 for *a*-stem kinship terms with honorific *-i*), *tiža* "heart," *f/paba* "mountain," *muža* "good order(?)" Divine names (some attested as elements of personal names only) include: *Šimiga* (beside *Šimige*), *Išhara*, *Tamgina* (*Damkina*), *Tilla*, *Naja*, *Ḫamanna*, *Ḫurra*, *Nuza*. For words with the suffix *-kka* see §4.4.2, 11.

4.4.3.3 o/u-stems

These stems mainly appear as names of non-Hurrian – in few cases also Hurrian – origin in the texts from Nuzi: *Marduku*, *Šelwuḫu*, *Kelžu*, *Kungu*, *Niru*, *Pendu*, *Šindu* (personal names), *Nullu* (country), *Nuzu* (city), *Šayu* (goddess[?], element in female personal names).

4.4.4 Athematic stems

Stems formed without a thematic vowel seem to occur more frequently in the earliest phase of the Hurrian language (mostly late third millennium BC). Some of the athematic stems later become thematic *i*- or *a*-stems: *šen* "brother" (cf. *šēna* and *-šenni*), *mad* "wisdom" (cf. *maḍi*), *adal* "strong" (also second millennium; seldom *adli*), *muḥ* divine name (cf. *muḥa*, *muḥni*), *Kaḥiari* name of the mountain *Ṭur^c Abdin* (cf. later *Kaḥiari*). Several names of gods, heroes, persons, and places are athematic: *Teššob* (cf. *Teššoba/i*), *Ḫebat* (cf. *Ḫeba*), *Kuḥuḡ* (cf. *Kuḥuḡa*, *Kuḥa*), *Nubadig*, *Taḥmiḥ*, *Šaluḥ*, *Šeriḥ* (cf. *Šeri*), *Ḫurriḥ* (cf. *Ḫurri*); *Gilgamiḥ*.

4.4.5 Honorifics

Some *a*-stems which denote human beings held in respect add a suffix *-i*: *alla=i* "lady, queen", *atta=i* "father"; **umma=i* "mother"? (attested only as personal name), *ušta=i* "hero." The name of the sun-god *Šimige* seems to be a contraction of *Šimiga=i* (cf. *Šimiga*); perhaps also *taḥḫel/taḡe* "man" from **taḥḫa=i* (cf. *taḡa* in personal names).

4.4.6 Derivational suffixes

These suffixes, which form either nouns or adjectives, follow the thematic vowel (and in rare cases also the possessive suffixes, for which see §4.4.8). Some of them (*-ni*, *-šše*) are identical in form with the noun-formation suffixes, but their position in the sequence of suffixes is different. In the case of thematic stems in *-i* the "derivational vowel" *-o*- or *-u*- replaces the thematic vowel, whereas stems in *-a* keep it. The derivational pattern has a parallel in the pronominal system which often shows an opposition between an absolutive ending in *-i* and oblique cases with *-o/u*- occurring before the case ending (see §§4.4.10.2 and 4.4.10.4). In very rare cases – apparently in old forms – the thematic vowel is not replaced by the derivational vowel. Moreover, some derivational suffixes follow a different pattern and do not replace *-i* by *-o/u*-.

The derivational suffixes are as follows:

1. *-ḡe*, *-ḫḡe*, adjectives of appurtenance: the form *-ḡe*, with the voiced initial consonant and used chiefly with geographical names, is treated in 1A–1E, *-ḫḡe* in 1F; both in 1G–1H.

1A. *i*-stems (*-ḡe*): *ḫurr/ḫurv=ō=ḡe* "Hurrian" (*Ḫurri*, **Hurvi*), *ḫatt=o=ḡe* "Hittite" (*Ḫatti*), *lupt=o=ḡe* "Luptian" (*Lupti* [a town]). When the word ends in *-ni*, *-li*, or *-ri*, the adjectives in *-ḡe* are commonly formed without the derivational vowel: *kibir=ḡe=n(a)=aḥ=a* (dat. pl., *Kibri* [a town]); *ḫamar=ḡe* "belonging to the *ḫamri*-sanctuary; *pabil=ḡ(e)=a* "in Babylonian" (from Akkadian *Bābili*); *bidin=ḡe* local form of a goddess (*Bidin* [a town]). The derivational vowel may, however, remain: *Ḫiri=ḡe* "wooded" (?) (name of a country, *ḫiri* "wood"?), *atta=šši=ḡe* "paternal property" (*attai* "father"), *ess=o=šši=ḡe* "?" (a kind of field).

1B. *a*-stems: *ankuwa=ḡe*, *ḫattarina=ḡe*, *šabinuwa=ḡe*, *tameninga=ḡe* (all based on names of cities), *alḫyḡ* (Ugaritic consonantal spelling) = **alaḫiya=ḡe* "Cyprian" (*Alaḫiya* "Cyprus").

1C. Athematic nouns: *tugriḥ=ḡe*, *mardaman=ḡe*, *igingallis=ḡe*, *aḫiḡis=ḡe* (all based on names of cities), *mugiḥ=ḡe* (*Mug/kiḥ*, name of a country).

1D. A special group of words: based on roots which are all attested in verbal use and which preserve *-i* (cf. §4.4.6, 4B): *pašš=i=ḡe* "consignment" (*pašš* "send"), *pa=i=ḡe* "ready for building (a house)" (*pa* "build"), *kunz=i=ḡe* "reverence" (?). (*kunz* "bow"), *un=i=ḡe* "offering" (?) (*un* "bring"), *na=i=ḡe* "pasture" (*nav* "graze").

1E. *Multiplicative numbers*: see §4.7.2, 4.

1F. *i-stems* (-*h₂he*): the form of the suffix with the initial doubled consonant, -*h₂he*, is used principally with *i-stems*: *hiyar*=*o*=*h₂he* "gold, golden" (*hiyari* "?"), *šiniber*=*o*=*h₂he* "of ivory" (**šinibēri* "ivory" < Akkadian *šinnipīri*), *ašt*=*o*=*h₂he* "female" (*ašti* "woman, wife"), *tur*=*o*=*h₂he* "male," *tiž*=*n*=*o*=*h₂he* "heart-shaped" (*tiž*=*ni* "heart"), *hažman*=*o*=*h₂he* "colored like the *hašmānu*-stone," *šimig*=*o*=*h₂he* "belonging to the sun-god" (name of a gate).

1G. *-ge/h₂he*-complexes: several suffix complexes seem to contain the suffix -*ge/h₂he*, such as the following: -*ašh₂*- (adjectives based on abstracts: *ašt*=*ašh₂* "female attributes," *aštašše* "womanliness," *ašti* "woman"); -*išh₂*- (*turišh₂i* "west," *turi* "low"); -*ušh₂*- (utensils: *ağr*=*o*=*šh₂i* "incense bowl," *ağri* "incense"); -*at₂h₂*- (mostly terms for household utensils: *kaz*=*o*/*ul*=*at₂h₂* a large bronze pot, from *kazi* "jar"?); -*it₂h₂*- (*pašš*=*it₂h₂i/e* "envoy," *pašš*-, "send"; *nir*=*an*=*it₂h₂* a kind of wood); -*o/ut₂h₂*- (*na₂h₂*=*o/ut₂h₂i* a seat).

1H. *Nouns of profession*: such nouns can be derived from adjectives of appurtenance by addition of the suffixes -*li* (see §4.4.2, 11) and -*ri*. For -*li* there are three patterns, presented here from least to most commonly occurring: the first (rare) preserves the suffix -*ge* unchanged: *šina*=*ge*/*i*=*l*- "crown prince; second quality" (*šina* "two"); the second shows the derivational vowel -*o/u*- before -*li*: *mardad*=*o*=*g*=*o*=*li* "carpet weaver" (from Akkadian *mardatu* "carpet"); and the third lacks the derivational vowel: *halz*=*o*=*g*(*e*)=*li* "district governor" (*halzi* "district"). The suffix -*ri* is seen, for example, in *am*=*om*=*i*=*h₂h₂(e)*=*o/u*=*ri* "administrator" (*am*- "see").

Derivational suffixes other than -*ge/h₂he*- are:

2. -*ni*, -*nni*, adjectives and nouns: the suffix -*ni* is found, for example, in *te*(*yi*)=*ō*=*n*=*ae* "widely" (**teyi*(?) "much"), *fağr*=*o*=*n*(*i*)=*ne*=*n* "beautifully" (*fağr*- "be beautiful"), *pic*=*o*=*n*(*i*)=*ne*=*n* "happily" (*pic*- "please"). Examples of -*nni* are: *mād*(*i*)=*o*=*nni* "wise" (*madi* "wisdom"), *attan*(*i*)=*o*=*nni* "father" (*attani* "father"), the personal names *Šenn*=*o*=*nni* (*šen*=*ni* "brother") and *Men*=*o*=*nni* (*men*=*ni* female relative), *haž*=*i*=*kk*=*o*=*nni* "deaf person" (*haž*- "hear," -*kk*- is a negative), and terms of profession like *urb*=*ar*=*i*=*nni* "butcher" (*urb*- "slaughter"), *fur*=*o*/*ull*=*i*=*nni* "diviner" (*fur*- "see").

3. -*ssi*, adjectives and nouns of suitability: *šen*(*a*)=*iffu*=*ssi* "suitable to my brother," *ašt*=*o*/*u*=*ss*- a garment (*ašti* "woman"), *pağ*=*o*/*u*=*ss*- a headgear (*pāgi*/*e* "head").

4. -*ži*/-*šše*:

4A. -*ži*, adjectives: *nīr*(*i*)=*o*/*u*=*ž*(*i*)=*ae* "well" (adverb in -*ae*, *nīr*- "be good"), *talāv*(*i*)=*o*=*ži* "great," *fağr*(*i*)=*ō*=*ži* "good," *šav*(*i*)=*ō*=*ži* "great."

4B. -*šše*/*i*, nouns: *itt*=*o*/*u*=*šše*- "garment" (*itt*- "clothe"), *suğr*=*o*/*u*=*šše* "meadow" (*suğri* "grass"), *nakk*=*o*/*u*=*šše* a military class (*nakk*- "release"). A special group of words in -*šše* are based on roots which are all attested in verbal use and which preserve the -*i*: *šar*=*i*=*šše* "desire" (*šar*- "wish, demand"), the personal name *Pašš*=*i*=*šše* "sending" (*pašš*- "send") (cf. §4.4.6, 1D).

5. -*bur*, negative: *mānn*=*ō*=*bur* "is not" (*manni* "is; he," see also §4.4.10.2 and §4.5.11); compare also *kuld*=*o*/*ubur* "?".

6. -*o/ubade*: For this morphologically unclear formant, see §4.4.2, 5.

4.4.7 Relational suffixes

The suffixes -*ne* (sg.) and -*na* (pl.) are anaphoric suffixes which are positioned between the noun and its case endings. They are incompatible with the possessive suffixes (except

perhaps in very rare, but still dubious cases) and they do not occur with names (except in a few cases of appellatives used as names like, e.g., *Kešše* “the one who sets (traps),” i.e., “hunter”). Singular *-ne* never occurs in the absolutive case, but *-na*, a plural marker, does. The two suffixes also precede most of the case endings which mark agreement of genitive modifiers with their head noun (*Suffixaufnahme*, see §5.2). Examples follow: *ērbī=ne=ž* “a dog” (ergative), *ōmin(i)=ne=ve allai* “the lady of the country,” *tažē=nē=va ed(i)=ī=da* “concerning the gift (dative)”; *paššitē=na* “the envoys,” *evren=n(a)=až=už* “the kings” (ergative), *ōmīn(i)=n(a)=až=a* “in the countries” (essive).

4.4.8 Possessive suffixes

These suffixes take the position after the noun-formation suffixes. They very seldom occur together with derivational suffixes; though in a few attested cases, they precede them. The pronominal element is clearly separate from the number suffix.

The possessive suffixes of the Mittani Letter – first, second and third person – are presented in (3):

(3)	Singular	Plural
1st	-iffə, -iffē-, -iffu-	-iff=až
2nd	-v/b/p	*-v=až (?)
3rd	-i (Hattuša: -ia/-iə)	-i=až

In a text from Hattuša, the second plural is attested twice: *ōlmi=šši* “your weapons”; *ede=ž=uda* “towards your body.”

4.4.9 Case and number suffixes

Hurrian is an ergative language. The agent of an action with explicit patient is marked as an ergative, and the patient as an absolutive. If the patient is not explicitly mentioned, the agent is encoded as an absolutive, as is the subject of an action or a state without implication of a patient:

- (4) A. *šēn(a)=iffu=šš(a)=ān ašti šār=ōž=a*
 “My brother (*šēna*, erg., with encl. pronoun 3rd per. sg.) has asked for a wife (*ašti*, abs.)”
 B. *šēn(a)=iffə pašš=ož=i*
 “My brother (*šēna*, abs.) has sent” (patient possible, but not mentioned)
 C. *tažē=n itt=ōš=t=a*
 “The gift (*taže*, abs.) has departed”
 D. *ēl(a)=iffə mănē=mmaman tupp=e*
 “My sister (*ēla*, abs.), she herself, is present”

In contexts not yet defined, a different pattern may replace the ergative one. In this instance the agent is encoded as an absolutive and the patient as an essive:

- (5) A. *fandarin(n)i=nā=ma ag=i=b neğern(i)=a*
 “And the cooks (*fandarinni*, abs.) took up breast meat (*neğerni*, ess.)”
 B. *el(i)=a fağr=o=ž(i)=a tan=d=i=b*
 “She gave (lit. ‘made’) a beautiful banquet (both essive)”

In total, nine or, in a wider definition, fourteen (see [7] below) cases have been identified so far. The plural is marked by three suffixes: (i) the relator plural *-na(-)* (see §4.4.7) for most noun cases (including the absolutive, conditioned, however, by the absence of a possessive pronoun): for example, *en(i)=na* "the gods," *en(i)=n(a)=až=e* "of the gods"; (ii) the plural suffix *-až-* (not in the absolutive after *-na*): *ēn(i)=iff=až=už* "our gods" (ergative), *en(i)=n(a)=až=už* "the gods" (ergative); (iii) the enclitic personal pronoun *-lla* (third-person plural; see §4.4.10.1), only in the absolutive: *en(i)=iffa=lla* "my gods."

The nine case suffixes of Hurrian are presented in (6). The ergative suffix *-ž* is absent before the enclitic personal pronouns except that of the third-person singular:

(6)	Singular	Plural
Absolutive	—	-na
	—	-lla
Ergative	-ž	-(na=)až=už
Genitive	-ve	-(na=)až=e (- (na=)aš=fe)
Dative	-va	-(na=)až=a (- (na=)aš=fa)
Directive	-da	-(na=)aš=ta
Comitative	-ra	-(na=)až=u=ra
Ablative-instrumental	-n(i)	?
Ablative	-dan(i)	-(na=)aš=tan
Directive	-ē	?

Conventionally, certain additional case suffixes have been identified. The absence of a plural in most instances and syntactic differences show their separate status.

(7)	Singular	Plural
Essive	-a	-až=a
Instrumental	-ae	—
Aequative	-ož	—
Associative	-nni	—
Associative-essive	-nn(i)=a	-až=o=nn(i)=a

4.4.10 Pronouns

In addition to the possessive suffixes of §4.4.8, Hurrian has personal and deictic pronoun suffixes:

4.4.10.1 Enclitic personal pronouns

These pronoun suffixes are restricted to the absolutive. They appear in two variants the distribution of which is not yet clear: a long form ending in *-a* (more frequent in the Mittani Letter) and a short form (more frequent in the texts from Hattuša and elsewhere):

(8)	Singular		Plural	
	long form	short form	long form	short form
1st	-tta	-d	-dilla	-dil
2nd	-mma	-m	-ffa	?
3rd	-nna	-n	-lla	-l

Only in the position after certain particles (see §4.6.1), the pronominal enclitic *-ma/e* is

used for the third-person singular. In the same position, *-lla* has an optional variant *-lle*.

4.4.10.2 Independent personal pronouns

Both these pronouns (except for the second person) and the deictic pronouns have an absolutive stem in *-e/i* and an oblique stem in *-o/u-*:

(9)	SINGULAR	<i>First</i>	<i>Second</i>	<i>Third</i>
	<i>Absolutive</i>	ište(=n)	fe	man(n)i
	<i>Ergative</i>	iž=až	fe=ž	manu=ž
	<i>Genitive</i>	šo=ve	fe=ve	—
	<i>Dative</i>	šo=va	fe=va	—
	<i>Directive</i>	šu=da	fe=u=da	—
	<i>Comitative</i>	šu=ra	—	manu=ra
	<i>Ablative</i>	—	—	manu=dan
	<i>Associative</i>	šo=nn(i)=a	—	—
	PLURAL	<i>First</i>	<i>Second</i>	<i>Third</i>
	<i>Absolutive</i>	šatti=(lla)	fe=lla	mane=l(la)
	<i>Ergative</i>	šiye=ž	fe=ž=už(?)	man=ž=ož
	<i>Dative</i>	—	fe=ž=a	man=z=a
	<i>Comitative</i>	—	—	man=ž=o/u=ra

In the plural, the genitive, directive, ablative, associative, and instrumental (also singular) are unattested.

4.4.10.3 Deictic pronouns

The system of deictic pronouns distinguishes between spatial and anaphoric deixis. There is a special *alternative* pronoun (the one-“the other”); only the anaphoric and the alternative pronouns make reference to the distinction “proximity versus distance”:

(10)			<i>Singular</i>	<i>Plural</i>
	<i>Demonstrative</i>	<i>absolutive</i>	anni	anni=l(la)
		<i>ablative</i>	annu=dan	
	<i>Anaphoric, proximity</i>	<i>absolutive</i>	andi	andi=lla
		<i>genitive</i>	andu=we	
		<i>dative</i>	andu=wa	
		<i>directive</i>	anduw=ē (?)	
	<i>Anaphoric, distance</i>	<i>absolutive</i>		ane=na/ani=lla
		<i>dative</i>	anu=wa	
		<i>ablative</i>	anū=dan	
	<i>Alternative, proximity</i>	<i>absolutive</i>	akki akki=lla	
		<i>ergative</i>	akku=ž	
		<i>ablative</i>	akku=dan	
	<i>Alternative, distance</i>	<i>absolutive</i>	agi	
		<i>dative</i>	agu=wa	
		<i>directive</i>	agu=da	

In addition, Old Hurrian shows an anaphoric resumptive pronoun *'alli*.

4.4.10.4 Interrogative and relative pronouns

This pronoun takes the form *ave-* "who." Attested is an ergative *ave=*ž=lla* "who... us?" (see §4.4.9).

4.5 Verbal morphology

Verbs seem to be marked for modes of action; some of the pertinent suffixes are only attested on verbal forms, whereas others modify the meaning of the root prior to the distinction of nominal or verbal inflection (see §4.3). The valence of a verb (transitive or intransitive; see §4.5.1) is indicated by the so-called class-markers. Valence may be modified either by changing the class-marker or by using a suffix which indicates intransitivity.

The verb in the Mittani Letter distinguishes three tenses (present, preterite, and future). Old Hurrian appears to distinguish aspect instead – it is not clear whether aspect is a category of the grammar of the Mittani Letter.

In ergative verb forms, three persons (first, second, third) and two numbers (singular, plural) are distinguished. The subject of nonergative forms in the Mittani Letter is not expressed by the verb form morphology, but only by a noun or by an enclitic pronoun (see §4.4.10.1) following the verb or any other constituent of the clause. For the subject suffixes of nonergative forms in Old Hurrian, see §4.5.9. Two negative suffixes, which are distinguished according to ergativity and nonergativity, are incorporated into the verbal form.

4.5.1 Valence

Valence (the number of noun phrases governed by the verb) is indicated by the vowels treated in §4.5.6. Some roots are attested in both transitive and intransitive use (*un-* "come"/"bring," *faž-* "enter"; *naḥḥ-* "sit down"/"set, place," *teğ-* "grow up/raise," *an-* "be pleased/please"). Normally, however, the root is attested in either transitive or intransitive usage. A change of valence appears to be marked by the suffix *-ol-*: *ḥic=ūḡ=i=vā=en* "may he not hurt [my heart]" *ḥic=ūḡ=ol=(a)=l=ē=tt(a)=ān* "I will grieve."

4.5.2 Modes of action

Several suffixes which immediately follow the root-complements (see §4.3) seem to mark modes of action:

1. The suffix *-il(l)-* marks the inchoative: *šid=ar=ill=ō=m* "he began to curse."
2. The function of the two suffixes *-ol-* and *-Všt-* (see §3.5.4, 4) is not yet clear; perhaps the first one marks duration and the second one result. In rare instances they may appear together: for example, *muž=ōl=ō=m* "he shaped [the goblet]"; *pa=ašt=ō=m* "he erected [a temple]"; *tav=ašt=ō=m* "he cast [a goblet]"; *teğ=ešt=a=b* "he grew up"; *ḥub=ušt=aw* "I shall break to pieces"; *til=ol=ōšt=aw* "I shall crush underfoot."
3. The rare verbal forms in *-uva* (*taž=ol=uva* "he made it shining") may also define a mode of action.

4.5.3 Undefined verbal suffixes

There are some more verbal suffixes occupying a position close to the root, the meaning of which has not yet been established: for example, *ešḥ-*, *-imd-*, *-upt-*, *-o/ušk-*, *-o/už-*, *upp-*.

4.5.4 Tenses

The tense suffixes are *-ož-* (*-ōž-* before *-t-*) for the preterite, and *-ēd-* (also *-ed-*) for the future. These suffixes have been explained as aspectual (perfective and imperfective respectively), but *-ož-* is never used for a complete action of the future, nor is *-ēd-* for a noncomplete action of the past. The present tense is morphologically unmarked.

4.5.5 The marker of a kind of direction(?) *-t-*

There is no agreement thus far concerning the function of the suffix *-t-* which follows the tense markers in intransitive verbs of movement, but seldom in ergative verbs.

4.5.6 “Class-markers” (suffixes of valence)

Old Hurrian distinguishes between three so-called “class-markers”: (i) *-a-*: one valence, intransitive, apparently only with verbs of motion; (ii) *-i-*: virtually two valences, but only one valence filled (see, however, the construction of §4.4.9 [5]), transitive-nonergative; (iii) *-o-*: two valences, ergative. In the Mittani Letter the forms in *-o-* are (nearly?) completely absent and have been replaced by forms with *-i-*. In ergative forms, *-i-* is not compatible with the tense suffixes *-ōž-* and *-ēd-*. Except in forms with the negative suffix *-ma* (see §4.5.7), it is also absent in the present tense before the personal suffixes of the first singular and plural.

4.5.7 Negation

Nonergative verb forms take the negative suffix *-kkV*, which follows the class-marker. The vowel agrees with the preceding vowel except before enclitic personal pronouns, where it changes to *-a-*. Intransitive negative forms replace the class-marker *-a-* by *-o-*, which, however, remains *-a-* before the enclitic personal pronouns. Ergative forms are negated by the suffix *-va-* (also found in both ergative and nonergative jussives; see §4.5.12.2) or *-ma-* and, especially in the dialect of the tablets from Hattuša, *-ud-* (also lexicalized: *sul=ud-*, *hemz=ud-* “loosen,” literally “untie”).

4.5.8 Ergative person suffixes

The following forms are attested in the Mittani Letter (and partially elsewhere):

(11)		<i>Singular</i>	<i>Plural</i>
	<i>1st</i>	<i>-aw</i>	<i>-aw=ž</i>
	<i>2nd</i>	<i>-o</i>	—
	<i>3rd</i>	<i>-a</i>	—

At Hattuša, the suffix of the second-person plural displays the ending *-āššo/ō* (written *-^oa-aš-šu(-u)*), which seems to invert that order of person and plural suffixes displayed in the first and third plural. A form of the third plural is attested at Hattuša: *-a=ž*.

4.5.9 Old Hurrian person suffixes

Old Hurrian has a suffix *-b* which seems to mark the third person (singular and plural) of intransitive and transitive-nonergative verbs. For ergative verbs, the suffix *-m* appears to mark the third-person singular of both agent and patient.

4.5.12.1 Imperative and cohortative

The imperative and the cohortative are formed by the root and the class-marker, optionally followed by an enclitic personal pronoun; the plural is marked by *-ž*. Both second- and third-person imperative forms occur, as well as a first plural cohortative:

(15)	Singular	Plural
1st		dilla...tād=ugār=i=ž "we wish to love each other" fağr=o=š=till(a)=ān "we wish to be friendly to each other"
2nd	un=a, un=a=mma "come!" ar=i 'give!' ḥaž=i=mma "listen!" kel=o, kel=o=m "be satisfied!"	sull=ud=i=ž "unbind!"
3rd	kud=o "let it be felled!" nakk=o=n(na) "let him/her be released!"	itk=o=ž "let [the temples] be purified!"

4.5.12.2 Jussive

The jussive expresses a request in the third person. Its suffix is *-en* (*-in*, in Hattuša also *-an* [personal correspondence from M. Giorgieri]), which in transitive forms follows the class-marker *-i-* or, in Old Hurrian, *-o-*. The final *-n* of the suffix could be a pronominal element (see §4.4.10.1), but it appears in forms of both the third singular and plural. The plural is marked by *-id-*. The negation of the jussive ("vetitive") is *-va-* (OH *-v(e)-*) after a vowel, *-ov-* after a consonant: *pašš=i=en* "he may send"; *tašp=o=in* "he shall destroy"; Hattuša *ar=i=an* "may he give"; *ḥa=i=en=i=lla* "may he take them"; *tād=ašt=id=en* "may they love us"; *itt=id=en* "may they go"; *ḥaž=āž=i=vā=en* "may he not listen [to them]"; *ur=ōv=en* "may he not exist"; *nakk=id=ōv=en* "may they not let/send."

4.5.12.3 Modal *-l-*

The modal suffix *-l-* (under undefined conditions apparently *-ll-*) combines with various mood forms and modifies their meaning in a way which, however, cannot always be well established.

4.5.12.4 Optative

The optative seems to be formed by the modal suffix *-l-* plus the jussive suffix (without its final *-n*). Third-person forms in texts from Hattuša often have the suffix *-ž* in a nonplural usage, presumably in an intensifying function. Forms from the Mittani Letter include: *ḥaž=i=l=e* "I wish to hear"; *kul=(*i=l=e* "I wish to say"; *ḥaž=āž=i=va=ll=i=lla* "I do not wish to hear them." From Hattuša come: *kad=i=l=e=ž* "may it speak," *kir=o=l=e=ž* "may it be relieved," *tag=o=l=e=ž* "may it be shining."

4.5.12.5 Potential

The potential is formed by the root and the suffix *-eva*: *ai...faž=ēva* "if [the enemy] invades"; ... *tta pic=ošt=ēva* "I would rejoice."

4.5.12.6 Conditional

The conditional is the potential modified by the modal suffix *-l-*; it is also used to express the contingency of an action: *kad=i=l=ēva* "[a word which somebody] might say"; *ḫill=o=l=eva* "he might say"; *ai=n ur=d=o=l=eva* "if it happened"; *ar=(*i=) r(<l)=eva=ž* "we are ready to give" (on the assimilation of *-l-*, see §3.5.4, 1)

4.5.12.7 Desiderative

The desiderative is a strong wish which may be modified (intensified?) by the modal suffix *-l-*: *itk=id=anni* "may they purify him/her"; *id=i=l=anni* "may he beat him."

4.5.12.8 Other possible modalities

In both the Mittani dialect and the dialect(s) of the Hattuša tablets, there occur roots with the suffix *-ai*, which Hittite scribes translated by subordinate clauses. Consider the following final (purpose) clauses: *faž=ai=n* "so that he may enter"; *itt=ai=ž=a=lla* "so that they may go"; *ḫaž=āž=ill=āi=n=i=lla* "so that he may (be ready to (? – inchoative)) hear them." Forms in *-ai* occur with a preceding *-l-* or *-m-*: *šid=i=l=āi* "so that he cursed [him]"; *naḫḫ=i=l=āi* "[he is someone whom his lord] appointed"; *fur=i=m=ai=n* "[whenever] he sees him"; *faž=o/u=m=ai* "when he entered"; *kunz=i=m=ai* "while he bows."

4.5.13 Verbal nouns

There is evidence of various Hurrian verbal nouns.

4.5.13.1 Infinitives

For the Hurrian infinitive, see §4.4.2, 16.

4.5.13.2 Nominalized verb forms

Finite verbs may be nominalized by the suffix *-šše* and treated like other nouns: *am=om=i=a=šše* "a dignitary" (ergative third singular); *ūr=i=ā=šše=na* "those which he desires"; *ar=ōž=aw=šše=nē=ve* "of that which I gave."

Several nominalized verb forms in the Mittani Letter contain an element *-mbū-* which has not yet been well defined (a state achieved as the result of an action?): *ur=i=mbū=šš(e)=o=ḫḫ(e)=a=mān* "and in the manner desired."

4.6 Particles

"Particle" will here be defined as a word which cannot take nominal or verbal suffixes but only enclitic particles (see §4.6.4) and pronouns.

4.6.1 Introductory particles

Certain particles introduce clauses: *adi* "so"; *ai* "if"; *alaže-* "whether"; *inna-* "when, as soon as"; *i/unu-* "as"; *panu-* "?"; *ia/e-* (a relative).

4.6.2 Adverbs

The following adverbial particles are identified: *anam(mi)* "in this manner"; *ḫenni* "now"; *kuro/u* "again, on the other hand"; *šukko* "once" (?); *tiššan* "very"; *undo* "now."

4.6.3 Interjections

The interjectory particles are *oia* "no!" and *au* "behold!"(?).

4.6.4 Enclitic particles

In the Mittani Letter, the enclitic particles are as follows: *-an*, *-mān* (connective for words and clauses); *-man* (emphatic, restrictive: "only"); *-mmaman* (emphatic(?)); *-nīn* (function unknown). Old Hurrian shows *-ma* (connective).

4.7 Numerals

Only the numerals 1 to 10, 13 or 30, 14(?), 17 or 70, 18 or 80, 10,000 and 30,000 are known.

4.7.1 Cardinals and ordinals

Ordinals are formed from cardinals by the suffix *-šše*, *-ži* (see §4.4.2, 15).

(16)	Cardinal	Ordinal
1	šukki, šuga(?)	?
2	šin(a)	šinzi
3	kig(a)	kiški (<*kik=ši)
4	tumni	tumušše, tumunzi
5	nariy(a)	narišše
6	šeže	?
7	šindi	šendešši
8	kira/i	?
9	tamri/a	?
10	eman	emanzi, emassi(?)
13/30	kigman(i)	
14(?)	šinašinda	
17/70	šindeman(i)	
18/80	kir(e)man	kirmanze
10,000	nubi	
30,000	kiga nubi	

4.7.2 Other numerals

Various other numeral formations are attested:

1. *Fractions*: *ša/eht- "one-half"; tumunzalli "one-quarter of a shekel."
2. *Collectives*: tumn=adi "four-spoked"; šež=adi "six-spoked"; eman=di "group of ten people."
3. *Distributives* (with instrumental suffix *-ae*): kig=ad(i)=ae "three each" (see §4.4.2, 1).
4. *Multiplicatives*: šukki "once" (see §4.4.6, 1E).
5. *Adverbs* (number with factitive, adjective, and essive suffixes): šug=am=ġ(e)=a "simple"; šin(a)=am=ġ(e)=a "twofold"; tamr=am=ġ(e)=a "ninefold"; eman=am=ġ(e)=a "tenfold"; šinz=o=ġ(e)=a "in the second place."
6. *Expressions of age* (only attested with Akkadian case ending): šin=arbu "two years old"; kig=arbu "three years old."

5. SYNTAX

5.1 Word order

In ergative clauses (see §4.4.9) the agent usually takes the initial position, followed by the patient and the verb (SOV):

- (17) pašš=īth(i)=iffu!=ž tīve andi kul=ōž=a
 "My messenger (*paššīthe*, erg.) said this word (*tīve*, abs.)"

Word order may be changed by topicalization, as seen in both the Mittani Letter (18A–B) and Old Hurrian (18C):

- (18) A. keb=ān=ož=āw=šše=na fur=ēd=ā=ll(a)=ān šēn(a)=iffu=ž
 "[The things] which I have sent (*keb=ān-*) my brother will see (*fur-*)"
 B. un=ā=l=an šēn(a)=iffu=wa
 "They do come (*un-*) to my brother"
 C. ḥa=i=en id(i)=ia=n nir=o/ubadi erāde=ne=ž
 "May the bird (*erāde*) take (*ḥa-*) the evil from his body (*idi*)"

Participants in the dative or directive may follow the verb, otherwise they are positioned between the ergative subject and the absolutive object. A modifier (including a genitive) may precede or follow its head.

Hurrian has postpositions, which may govern a preceding dative or genitive. The following are found with a preceding dative: *ed=ī=da* (directive of *edi* "person, body," with a third-person singular possessive pronoun) "with reference to, concerning"; *e/ig=ī=da* "within"; *fur=ī=da* (*furi* "eye") "with regard to"; *ā(i)=ī=da* "in the presence of" (at Hattuša occurs also the preposition *ābi*). Governing a preceding genitive are: *āi=ē* (directive) "in front of"; *ed=ī=ē* "because of, about."

5.2 Agreement

A modifier agrees with its head. The case endings copied from the head are preceded by *-ne-* or *-na-* (see §4.4.7). This also applies to genitive modifiers (*Suffixaufnahme*):

- (19) A. šēn(a)=iffu=we=nē=ž ašt(i)=i=ž
 "My brother's (*-we* gen.) wife (*ašti*)"
 B. šēn(a)=iffu=we=nē=va torub(i)=ī=va
 "To my brother's enemy (*torubi* 'enemy,' *-va* dat.)"
 C. en(i)=n(a)=āž=(v)c=ne=da šarri=ne=da
 "To the king (*šarri*, *-da* dir.) of the gods (*en(i)=na*)"

Likewise, nominalized ergative verbs are constructed as modifiers which agree with their head. In this case, the head is always the patient of the nominalized verb, regardless of its case form:

- (20) A. šēn(a)=iffu=ll(a)=ān ūr=i=ā=šše=na tivē=na
 "The things (*tīve*) which my brother desires (*ūr-*)"
 B. tuppe niḡār(i)=rē=ve ar=ōž=aw=šše=nē=ve
 "The tablet (*tuppe*) of the dowry which I have given (*ar-*)"

5.3 Coordinate and subordinate clauses

There seem to be no special verbal forms for subordinate clauses except the verbal nouns mentioned in §4.5.13.1. Particles occurring in initial position of temporal, conditional, comparative, and other clauses have been cited in §4.6.1.

In relative clauses introduced by the particle *īa-*, *īe-*, the head of the relative clause is incorporated within the clause; the verb is nominalized and stands in agreement with the head. The main clause refers to the head of the relative clause by an anaphoric pronoun:

- (21) [[īa=llā=nīn šēn(a)=iffu=ž...tivē=na tān=ōž=ā=šše=na]_{REL.CL.}...andi=ll(a)=ān Šimīge=ne=ž ar=ēd=a šēn(a)=iffu=wa]
 “The things which my brother has done, those the Sun-god will give to my brother”

In rare cases the verb of the main clause may be incorporated:

- (22) [ia=mē=nīn ed(i)=iffə pal=āw [šēn(a)=iffu=ž...tād=i=ā=šš(e)=a]_{REL. CL.}]
 “I know (*pal-*, erg.) that [my brother loves (*tād-*, erg., nominalized, essive) it], my person (*ed(i)=iffə*)”

6. LEXICON

Hurrian is still only very incompletely known, especially as far as the lexicon is concerned. Since Hurrians had been in contact with the peoples of the northeastern parts of the Fertile Crescent since at least the last quarter of the third millennium BC (and presumably much earlier), one should expect a considerable stock of Sumerian as well as Akkadian and other Semitic loanwords. There are indeed some words borrowed from Sumerian in the third millennium, like *en=dan* “ruler” (from *en*); other possible Sumerian loans are disputed (*natḫi* “bed,” Sumerian *ná* “bed”). Akkadian loanwords are numerous, especially in texts from the Late Bronze Age. They reflect an extended usage of Akkadian as a second language, or at least as the vernacular of written communication and documentation; examples include: *šarri* “(divine) king” (from *šarru* “king”); *šukkalli* “vizier” (from *sukkallu*); *tupšarri* “scribe” (from *tupšarru*); *tamgarašše* “profit” (from *tamkāru* “merchant”); *salamži* “statue” (from *šalmu*); *ḫassissi* “ear” (from *ḫasīsu*); *arni* “guilt” (from *arnu*).

The tribes who established the Mittani dynasty spoke an archaic form of Indo-Aryan, which left some traces in Hurrian: *mariyanni* “chariot owner” (cf. Sanskrit *mārya-* “young man”); *pabro/unni*, *paridanni*, *pingaranni* colors of horses (cf. *babhrú-* “brown,” *palitá-* “grey,” *piṅgalá-* “reddish”).

Though Hurrian played an important role in Anatolia as a language of learning and ritual, it appears not to have borrowed from the Anatolian Indo-European languages (Hittite, Luwian, etc.) – an appearance perhaps connected with the fact that the Hurrian tablets chiefly preserve texts of Hurrian traditions. There is little doubt, however, that Hurrian was not only a literary language, but was spoken in a court and temple milieu at least in the fourteenth century BC. This is revealed by the occurrence of adjectives based on Anatolian place names: *ḫattoḡe* (“belonging to Hatti”), *šabinuvaḡe* (“belonging to Šapinuwa”), and so forth.

Bibliography

There is no Hurrian grammar or dictionary which presents the present state of our knowledge of the language. The following monographs have to be corrected and supplemented in the light of the research literature of the last decades:

- Bush, F. W. 1964. *A Grammar of the Hurrian Language*. Dissertation, Brandeis University. Ann Arbor: University Microfilms.
- Diakonoff, I. M. 1971. *Hurrisch und Urartäisch*. Munich: Kitzinger.
- Khačikjan, M. L. 1985. *Churritskij i urartskij jazyki*. Jerevan: AN Armanskoj SSR.
- Laroche, E. 1980. *Glossaire de la langue hourrite*. Paris: Klincksieck.
- Speiser, E. A. 1941. *Introduction to Hurrian*. New Haven: American Schools of Oriental Research.
- Wilhelm, G. 1989. *The Hurrians*. Warminster: Aris and Phillips.

After the completion of this grammatical outline (spring 1999), two grammars of Hurrian have appeared:

- Giogieri, M. 2000. "Schizzo grammaticale della lingua hurrica", *La parola del passato* 55: 71–277.
- Wegner, I. 2000. *Hurritisch. Eine Einführung*, Wiesbaden: Harrassowitz.

Sanskrit

STEPHANIE W. JAMISON

1. HISTORICAL AND CULTURAL CONTEXTS

Sanskrit is an Indo-European language, a member of the Indo-Aryan branch of the Indo-Iranian subgroup of that family. It is chronologically and in terms of linguistic development the “oldest” Indo-Aryan language and consequently often referred to as *Old Indic* (Altindisch) or *Old Indo-Aryan*; its descendants include a range of linguistic varieties classified under the rubric Middle Indic (or Prākṛit, see Ch. 27), as well as the Modern Indic (New Indo-Aryan) languages spoken today, such as Hindi, Gujarati, Bengali. It is not related genetically to the Dravidian languages of South India, such as Tamil and Telugu.

The oldest form of Sanskrit is so-called *Vedic Sanskrit*, the language of the four collections of liturgical texts known as the Vedas and of the early exegetical literature on these texts. The oldest Veda is the *Ṛgveda* (*Rig-veda*), a compilation of 1,028 hymns which took shape around 1500 BC in northwest India, though the composition and collection of hymns clearly occupied several centuries. In language, style, and phraseology the *Ṛgveda* resembles the earliest texts of its closest linguistic relative, the *Gāthās* attributed to the prophet Zarathustra, composed in Old Avestan (see Ch. 29).

Though the composition of Vedic texts can be dated with fair confidence to the period of c. 1500–500 BC, direct records of them are only found several millennia later. The “texts” were transmitted orally, with minimal alteration, and even after they were also committed to writing, the manuscripts were perishable and less reliable than the oral tradition.

Through the approximately thousand years of Vedic textual composition, the language shows gradual changes, especially in the loss of certain grammatical categories and the reduction of variant forms. Around 500 BC the Sanskrit then current among cultivated speakers received a magnificent description by the grammarian Pāṇini in his treatise, the *Aṣṭādhyāyī* (“[Work] consisting of eight chapters”), whose level of detail and theoretical sophistication has not been equaled to this day.

Pāṇini inadvertently froze the language in this particular form forever. What was composed as a *descriptive* grammar (though descriptive of a geographically and socio-culturally limited speech form, not the speech of the whole society) became a *prescriptive* grammar of a learned language. All subsequent Sanskrit follows, or attempts to follow, the rules of Pāṇini. Though there are systematic variations in later texts, these are essentially stylistic and distributed according to textual genre. The language of the great epics, the *Mahābhārata* and the *Rāmāyaṇa*, deviates somewhat from the Pāṇinian norm and is therefore sometimes distinguished as *Epic Sanskrit*; it displays some Middle Indic tendencies. Inscriptional Sanskrit also commonly shows nonsanctioned forms. Despite these minor exceptions, Sanskrit no longer had a history in the accepted linguistic sense of this term – even though the greater

part of its literature remained to be composed. The great flourishing of Sanskrit literary production lasted through the first millennium AD.

The language as fixed by Pāṇini is commonly known as *Classical Sanskrit*, or *Sanskrit* proper. Indeed, the term *saṃskṛta* means “perfected” and refers to the language generated according to Pāṇini’s rules, as opposed to the vernacular *Prākṛits*, from *prākṛta* “natural, unrefined.” Strictly speaking, the pre-Pāṇinian language of the Vedic texts is not “Sanskrit,” and is sometimes called simply *Vedic*, rather than Vedic Sanskrit. In this work, however, *Sanskrit* will denote all varieties of Old Indic.

2. WRITING SYSTEM

The earliest Sanskrit texts were composed and transmitted orally, not written down for centuries after their first “attestation.” Indeed, the first documentary evidence of Indo-Aryan languages in the Indian subcontinent comes not from Old Indic but Middle Indic: the inscriptions of the ruler Aśoka in the third century BC (see Ch. 27, §1.1). The first direct attestation of Sanskrit comes from around the beginning of the present era. The first extensive inscription is that of the ruler Rudradāman c. AD 150 at Girnar in western India; the first extant manuscripts, found in central Asia, date from about the same period.

The writing system found in most of the early inscriptions is *Brāhmī* (another, less widespread system, *Kharoṣṭhī*, an adaptation of Aramaic, is found in the northwest, already in the Aśokan edicts). *Brāhmī* seems to have been adapted from a Semitic writing system, though the exact details are unclear, as is the date of its introduction into India, a subject of much controversy. *Brāhmī* is the ancestor of most of the writing systems used in India.

Until the advent of printing and the regular publication of Sanskrit texts, Sanskrit manuscripts were written in various local scripts. Now Sanskrit is almost exclusively printed in a script known as *Nāgarī* or *Devanāgarī*, a medieval offshoot of *Brāhmī*, and perfectly adapted to the writing of Sanskrit, with a one-to-one correspondence between sound and symbol. The conventional transcription of *Devanāgarī* into Roman characters was established finally at the Tenth Congress of Orientalists, 1894. Transliterations in works published before often show deviations from the modern norm.

The system can be considered a modified or pseudo-syllabary in that each consonantal symbol represents a consonant with following short *a*-vowel (the commonest vowel in the language), for example, क = *ka*, ख = *kha*, ग = *ga*, घ = *gha* (not *k*, *kh*, *g*, *gh*); see Table 26.1. However, unlike “pure” syllabaries, a different symbol is not necessary to represent consonants followed by other vowels (e.g., *kā*, *ki*, *kī*, etc.). Instead, a set of universally applicable diacritics can be used to cancel the inherent short *a* and substitute a different following vowel: thus, क = *kā*, कि = *ki*, कु = *ku*, and so forth. There are also separate signs for independent vowels, for example, अ = *a*, ए = *e*.

Another drawback of some syllabaries, the inability to represent consonant clusters unambiguously, is overcome by the system of ligatures. Portions of each consonant in a cluster are combined into a single conventional sign, for example, त (ta) + क (ka) = त्क (tka). Final consonants can also be represented, by a stroke (*virāma*) under the sign, which cancels the short *a*: thus त = *ta*, but त् = *t*. Thus, the system combines the flexibility of an alphabet with some of the spatial economy of a syllabary.

Devanāgarī writing of Sanskrit lacks word divisions. Each linguistic string, regardless of morphosyntactic structure, is treated as a sequence of syllables (*akṣaras*) consisting of onset

Table 26.1 The Devanāgarī script

Vowel symbols

a	ā	i	ī	u	ū	ṛ	ṝ	l
अ	आ	इ	ई	उ	ऊ	ऋ	ॠ	ऌ
e	ai	o	au					
ए	ऐ	ओ	औ					

Consonant + vowel symbols

ka	kha	ga	gha	na
क	ख	ग	घ	ङ
ca	cha	ja	jha	ña
च	छ	ज	झ	ञ
ṭa	ṭha	ḍa	ḍha	ṇa
ट	ठ	ड	ढ	ण
ta	tha	da	dha	na
त	थ	द	ध	न
pa	pha	ba	bha	ma
प	फ	ब	भ	म
ya	ra	la	va	
य	र	ल	व	
śa	ṣa	sa	ha	
श	ष	स	ह	

Sample vowel diacritics

kā	kī	kṛ	ku	kū
का	कि	कृ	कु	कू
ke	kai	ko	kau	
के	कै	को	कौ	

consonant(s) (if present) plus vowel. Thus, a string like *tad etad rūpam*, with word divisions as given in transliteration, would obligatorily appear in Devanāgarī as *ta de ta drū pa m* त दे त द्रूपम् (though without spaces between the characters).

3. PHONOLOGY

3.1 Diachronic overview

From the point of view of reconstructed Proto-Indo-European, the most important phonological development in Sanskrit (and indeed in Indo-Iranian) is vowel-merger: short *e, *o, and *a all merge as a; long *ē, *ō, *ā (and short *o under certain conditions) merge as ā. Since much of Proto-Indo-European morphology was based on alternations of vowels with *e-timbre and those with *o-timbre (qualitative ablaut), these mergers had major effects on the morphological system.

On the other hand, Sanskrit maintained the Proto-Indo-European consonantal system with some fidelity, only enlarging its inventory. The three series of stops – voiceless (T), voiced (D), and voiced aspirated (Dh) – traditionally reconstructed remain in Sanskrit, and

a fourth was added, voiceless aspirated (Th). As in other satem languages the labiovelars merged with the plain velars. There was secondary palatalization of the resulting segments, reflected in thoroughgoing synchronic alternations within Sanskrit (see §3.4.2.2). Otherwise, the inventory of places of articulation was increased by the creation of a series of retroflex dental stops. For the comparatist an especially important retention in Sanskrit is the preservation of *y, *w, and *s intervocalically, thus avoiding the loss of morphological clarity attendant on vowel contraction that bedevils the historical linguist in languages like Greek.

3.2 Vowels

The cardinal vowels *i*, *u*, *a* distinguish length; in addition, short *a* is a closer vowel than *ā*, equivalent to schwa. The mid vowels *ē* and *ō*, as monophthongizations of the Indo-Iranian diphthongs **ai* and **au* (preserved in Iranian), are inherently long and are so marked in the phonological sections of this work, though they are not usually so transcribed. The true diphthongs *āi* and *āu* (usually now transcribed simply *ai* and *au*) also count as long. The vocalic liquid *ṛ* represents a merger of PIE (Proto-Indo-European) **r* and **l*. However, long *ṝ* is an invention of the system and found in a few analogically generated morphological categories; PIE **r̄* has different, biphonemic outcomes in Sanskrit, as we will see. Vocalic *ḷ* is even more limited, found in only one morpheme.

(1) Sanskrit vowel phonemes

monophthongs:	<i>i</i> / <i>ī</i>	<i>u</i> / <i>ū</i>	diphthongs:	<i>āi</i>	<i>āu</i>
	<i>ē</i>	<i>ō</i>			
	<i>a</i>		vocalic liquids:	<i>ṛ</i> / <i>ṝ</i>	<i>ḷ</i>
	<i>ā</i>				

3.3 Consonants

The consonantal inventory of Sanskrit is presented in Table 26.2:

Table 26.2 The consonantal phonemes of Sanskrit

Manner of articulation	Place of articulation					
	Labial	Dental	Retroflex	Palatal	Velar	Glottal
Stops and affricates						
Voiceless	p	t	ṭ	c	k	
Voiceless aspirated	ph	th	ṭh	ch	kh	
Voiced	b	d	ḍ	j	g	
Voiced aspirated	bh	dh	ḍh	jh	gh	
Nasals	m	n	ṇ	ṇ̄	ṇ̄	
	ṁ <i>anusvāra</i> (see below)					
Fricatives						
Voiceless		s	ṣ	ś		ḥ <i>visarga</i>
Voiced						h
Liquids		l	r			
Glides	v			y		

The apparent symmetry of this consonantal system conceals some failures of parallelism in distribution, often the results of historical changes:

1. The voiceless aspirated series is an addition to the system and significantly rarer than the other three. It is often found in etymologically obscure words.
2. The retroflex sibilant \mathring{s} is the automatic product of dental s following i , u , r , and k (mnemonically “ruki”), a process also found not only in Iranian but in part in Balto-Slavic.
3. The series of retroflex stops was a creation of Indic, in most cases as a conditioned result of regressive assimilation to rukified \mathring{s} and therefore distributionally limited; in particular, initial retroflexes are almost never found. The retroflex nasal is ordinarily the automatic product of dental nasal when the word contains a preceding r (subject to some conditions). Thus, all the retroflexes are in origin conditioned alternants of dentals, though from the beginning of the language they have a qualified independence.
4. The palatals are affricates, not stops. In the palatal row the voiced aspirate jh is a new and extremely rare phoneme; the phoneme patterning with the palatals as the voiced aspirate for morphophonemic purposes is glottal h (see §3.4.2.1).
5. The palatal nasal is a conditioned variant of n occurring next to palatal obstruents; the velar nasal is also ordinarily a conditioned product of n , found before velar stops, but further phonological developments (loss of final or cluster-internal velar stop) can allow the velar nasal an independent if marginal existence. *Anusvāra* is a conditioned alternant of postvocalic nasals, under certain sandhi conditions.
6. *Visarga* is a word-final (sometimes morpheme-final) conditioned alternant of s and r under certain sandhi conditions.
7. The glides and liquids regularly alternate with vowels: $i \approx y$; $u \approx v$ ($[w]$); $r \approx \mathring{r}$; $l \approx \mathring{l}$ (under conditions discussed below).

3.4 Phonological alternations

Sanskrit is characterized by a pervasive series of phonological alternations occurring on several different linguistic levels and displaying varying degrees of transparency. We begin with the most transparent.

3.4.1 External sandhi

The surface form of any linguistic string is subject to phonological rules of combination (*sandhi* or “putting together”). In other words, phenomena of the English *gonna* (from *going* + *to*) type apply to any two words in contact within a sentence, and even between sentences in a discourse. Most sandhi rules involve regressive assimilation, especially in voicing: for example, (with underlying *tad*) *tad bhavati* but *tat phalam*. Assimilation in manner of articulation is also met with (e.g., *tan manas*). Like vowels coalesce into a single long vowel (e.g., *vada* + *agne* \Rightarrow *vadāgne*), and unlike vowels undergo diphthongization or glide-formation (e.g., *vada oṣadhe* \Rightarrow *vadauṣadhe*; *asti agniḥ* \Rightarrow *asty agniḥ*). Despite the simplicity of the principles, the details of sandhi rules are sometimes opaque. For example, though the change of final *-as* to *-o* before voiced sounds historically involves regressive voicing assimilation, this process is not synchronically transparent. The rules of external sandhi ordinarily apply also at compound seams, and many but not all of the same rules at morpheme boundaries.

External sandhi in Vedic is more variable than in Classical Sanskrit, not only in the form of the rules but also in their application (or nonapplication). Sandhi in Middle Indic occurs only under conditions of close syntactic nexus. Given these facts, it seems likely that the pervasive system of obligatory sandhi characteristic of Classical Sanskrit involved an artificial imposition of an originally more flexible set of processes linking words within syntactically defined phrases.

3.4.2 Internal consonantal alternations

The rest of this section presupposes the concept of the *root* and the canonical structure of the Sanskrit word presented in §4.1.

3.4.2.1 Voicing and aspiration

The voiceless, voiced, and voiced aspirated obstruents of a positional series regularly alternate with each other ($p \approx b \approx bh$; $t \approx d \approx dh$, etc.; note, however, $c \approx j \approx h$), such that, for example, a morpheme with an underlying voiced aspirate final may show alternants with all three stops under differing internal sandhi conditions: thus, \sqrt{buddh} “be aware” – *buddh-yate*, *bud-dha-*, *bhot-syate*.

Clusters containing unaspirated stops show regressive assimilation (e.g., *chit-ti-* from $*chid + ti-$). But in those containing voiced aspirates the resulting cluster is both voiced and aspirated whatever the position of the aspirate in the underlying cluster (hence *buddha-* from *buddh-ta-*) – the change known as *Bartholomae's Law*. In summary,

$$\begin{array}{lll}
 (2) \quad A. \ T + T & B. \ T + D & C. \ Dh + T \\
 & \Rightarrow T-T & \Rightarrow D-D \\
 D + T & D + D & Dh + D \Rightarrow D-Dh \\
 & & Dh + Dh
 \end{array}$$

Before *s* all stops become voiceless; hence *bhot-syate* above. This same form illustrates another, sporadic alternation: when roots with underlying final aspirates lose that aspiration, the initial consonant often acquires aspiration (hence *bhot-syate*, but *buddh-yate*). This represents a reconfiguring or reversal of a historical development – *Grassmann's Law*, whereby di-aspirate roots dissimilated the first aspirated stop.

3.4.2.2 Velars and palatals

The velar series (k, g, gh) regularly alternates with the palatal series (c, j, h). In particular, velar-initial roots reduplicate with palatals (e.g., \sqrt{kr} : *ca-kāra*; \sqrt{gam} : *ja-gāma*); and preobstruent velars alternate with palatals in other phonological positions (e.g., \sqrt{muc} : *muk-ta*, but *muc-yate*). This alternation is the historical result of a pan-Indo-Iranian palatalization of velars by following front vowels (and y), the conditioning of which was obscured by the subsequent merger of $*e$ with $*a$ and $*o$ noted above.

3.4.2.3 Palatals and retroflexes

The structural position of the palatal series was further complicated by a different merger. Though Sanskrit c can only be the product of an old palatalized velar ($*k^{(w)}e$, etc.), both j and h have two sources: (i) not only palatalized velars ($*g^{(w)}e$, $*g^{(w)h}e$, etc.); (ii) but also PIE palatal stops ($*ǵ$ and $*ǵ^h$), whose voiceless equivalent ($*ǵ$) yields Sanskrit $ś$. These underlying palatals enter into a set of synchronic alternations different from that of the old velars: palatals followed by dentals produce a retroflex cluster, for example, \sqrt{srj} “emit”: $srj + ta \Rightarrow srś-ṭa$. Thus, though the phonetic inventory of the language contains only a single palatal series, morphological alternations define two morphophonemically distinct series: (3A) $ś, j, h$; and (3B) c, j, h .

- (3) A. *palatals* (\approx *retroflexes*) B. *palatalized velars* (\approx *velars*)
 ś (e.g., viś : viṣ-ṭa) c (e.g., muc : muk-ta)
 j (e.g., sṛj : sṛṣ-ṭa) j (e.g., bhaj : bhak-ta)
 h (e.g., ruh : rū-ḍha) h (e.g., snih : snig-dha)

(By Bartholomae's Law, a compensatorily lengthened vowel plus retroflex $\dot{d}h$ is the regular outcome in *rūḍha*.)

The distinction between these two series is neutralized before *s*, where both series (and all three manners) show *k*: for example, both *ruruk-ṣati* (\sqrt{ruh}) and *sisnik-ṣati* (\sqrt{snih}).

3.4.3 Internal vocalic alternations

The Sanskrit morphological system is pervaded by vocalic alternations, conveniently considered as the "strengthening" of an underlying vocalic element by two successive additions of the vowel *a* ($V / \dot{a} + V / a + aV$). The preconsonantal versions of these strengthenings are known by the indigenous terms *guṇa* and *ṽddhi*, but it is useful to consider these in conjunction with their prevocalic alternants as well. In terms familiar from Indo-European descriptive grammars, the unstrengthened state corresponds to *zero-grade*, *guṇa* to *full-* (or *normal-*) grade, and *ṽddhi* to *extended-* (or *lengthened-*) grade. Though Proto-Indo-European *qualitative* ablaut essentially disappeared in Indo-Iranian with the merger of **e* and **o*, *quantitative* ablaut is transparently continued by the Sanskrit system of vowel strengthening. Alternations between zero-grade and full-grade are prominent in the morphological system; *ṽddhi* is especially important in the derivation of adjectives of origin and appurtenance (*ṽddhi derivatives*).

The alternations between consonantal and vocalic versions of glides and liquids are also relevant here, and the system is in fact clearest with these segments, especially *ṛ*, where the successive additions of *a* are easily discerned (N.B. for ease of exposition, *ī* and *ū* are not included here, but will be discussed below):

(4)	<i>Zero-grade</i>		<i>Full-grade</i>		<i>Extended-grade</i>	
	PreC.	PreV.	PreC. (<i>guṇa</i>)	PreV.	PreC. (<i>ṽddhi</i>)	PreV.
ṛ	ṛ	r	ar	ar	ār	ār
i	i	y	ē	ay	āi	āy
u	u	v	ō	av	āu	āv
a	a		a		ā	
ā	ā		ā		ā	

As can be seen, with the simple vowels *a* and *ā*, the progressive addition of *a* is not so clear; moreover, prevocalic position is subject to complications.

Though Sanskrit does not have surface syllabic nasals (**ṁ* and **ṇ*) as reconstructed for Proto-Indo-European, the parallelism of morphological alternations compels us to posit such underlying vowels, which fit into the vowel gradation system as follows:

(5)	<i>Zero-grade</i>		<i>Full-grade</i>		<i>Extended-grade</i>	
	PreC.	PreV.	PreC. (<i>guṇa</i>)	PreV.	PreC. (<i>ṽddhi</i>)	PreV.
<i>*ṁ</i>	a	m	am	am	ām	ām
<i>*ṇ</i>	a	n	an	an	ān	ān

The following chart gives an example of a root with each vocalism (save for the *a*-vowels), with representative forms from the various categories:

(6)		<i>ṛ</i> kṛ “do, make”	<i>i</i> ji “conquer”	<i>u</i> su “press”	* <i>ṛ</i> gam “go”	* <i>ṛ</i> han “smash”
Zero:	PreC.	kṛ-ta	ji-ta	su-ta	ga-ta	ha-ta
	PreV.	cakr-ur	jigy-ur	suṣv-āna	jagm-ur	jaghn-ur
Full:	PreC.	kar-tum	jē-tum	sō-tum	gan-tum	han-tum
	PreV.	akar-am	jay-ati	asūṣav-at	agam-am	ahan-am
Ext.:	PreC.	kār-ya	ajāi-s	asāu-ṣīt	—	—
	PreV.	cakār-a	ajāy-i	asāv-i	jagām-a	jaghān-a

3.5 Syllable structure and phonotaxis

There are few constraints on syllable structure. Syllables may both begin and end with vowels, single consonants, or consonant clusters; and internal vowels may be of any weight, even before coda consonants. In Vedic, however, some traces of phonological processes (*Sievers-Edgerton Law*) seemingly function to avoid overlong syllables: some suffixes containing *y* or *v* must be read as *iy* and *uv* after heavy syllables, but *y* and *v* after light syllables. But this is a morphologically limited phenomenon, not a pervasive phonological rule.

There are constraints on word-final consonants, which apply before external sandhi rules operate. Final clusters are not allowed (though monomorphemic *r*+ obstruent is rarely retained), and certain classes of sounds, such as aspirates and palatals, are not permitted finally.

3.6 Accent

Vedic Sanskrit has a pitch accent system, described also by Pāṇini, but accent has disappeared in Classical Sanskrit. The Vedic accent can fall anywhere in the word and, as it is not phonologically predictable, the position of the accent often conveys morphological and syntactic information.

Most Vedic words possess one accent. A few loosely bound compounds keep accent on both members, and a number of linguistic forms lack accent: some particles, some pronouns, and, most interestingly, noninitial vocatives and noninitial finite verbs in main (but not subordinate) clauses.

For ease of exposition, accent will in many instances not be marked in the ensuing discussions.

3.7 Diachronic developments

As in most early Indo-European languages, the loss of the so-called *laryngeal* consonants (cover-symbol **H*) of Proto-Indo-European had major effects on Sanskrit phonology and morphology. The phonological alternations originally caused by these segments have been morphologized in various ways, especially visible in the variant forms of roots.

1. *seṭ* vs. *aniṭ* roots: In many obstruent-final roots, an *i* (from vocalized **H*) surfaces in preconsonantal position, with no counterpart prevocally. Such roots are known as *seṭ* (“with an *i*”), and contrast with apparently parallel *aniṭ* (“without an *i*”) roots. Compare examples of identical morphological categories:

	<i>seṭ:</i>	<i>aniṭ:</i>
	pat "fly"	cit "think"
PreC.	pati-ta	cit-ta
PreV.	pat-ati	cet-ati

Because the distinction is neutralized in prevocalic position and because the interposition of the *i* helps to avoid the often awkward sandhi of consonant clusters, this *i* spreads beyond its original historical boundaries. Indeed, many suffixes and endings are reinterpreted as having an initial *i* (or at least an alternate form with initial *i*).

2. *Roots in *ṚH*: Sonorants (or resonants; i.e., *i*, *u*, *ṛ*, (*ḷ*); *ṛ, *ṛ) followed by a laryngeal in Proto-Indo-European yield so-called *long sonorants*, having root-final alternation patterns as follows:

(8)	<u>Zero-grade</u>		<u>Full-grade</u>		<u>Extended-grade</u>	
	PreC.	PreV.	PreC. (<i>guṇa</i>)	PreV.	PreC. (<i>vṛddhi</i>)	PreV.
*ṛH	īṛ/ūr	ir/ur	ari	ar	āri	ār
*iH	ī	(i)y	ayi (> ē)	ay	āyi	āy
*uH	ū	(u)v	avi	av	āvi	āv
*ṁH	ām	(a)m	ami	am	āmi	ām
*ṇH	ā	(a)n	ani	an	āni	ān

Consider the following examples:

		*ṛH tṛ "cross"	*iH nī "lead"	*uH bhū "become"	*ṁH kram "stride"	*ṇH jan "be born"
Zero:	PreC.	tīr-ṇa	nī-ta	bhū-ta	krām-ta	jā-ta
	PreV.	tīr-ati	nin(i)y-ur	bhuv-āni	cakram-ur	jajñ-ur
Full:	PreC.	tari-ṣyati	nayī-tum	bhavi-tum	krami-ṣyati	jani-tum
	PreV.	tar-ati	nay-ati	bhav-ati	kram-ate	jan-ati

The distribution of *īṛ/ūr* and *ir/ur* forms in *ṛH roots was originally conditioned by the quality of the preceding consonant, with *u*-forms following labials (e.g., √pṛ "fill," with *pūrṇa*).

3. *Roots in ā*: Such roots show an extremely anomalous set of alternations in comparison with the patterns set by other root types. As was first recognized by F. de Saussure in the 1870s, the anomalies can be explained by positing the same structure and alternations as in *seṭ* roots; in other words, by rewriting (in modern terms) *ā* as *VH and its unstrengthened form as *H, yielding *i* before consonant and zero before vowel:

(10)	<u>Zero-grade</u>		<u>Full-grade</u>
	PreC.	PreV.	PreC. (<i>guṇa</i>)
(*VH) > ā	(*H) > i	ø	ā
for example, sthā "stand"	sthi-ta	tasth-ur	asthā-t

4. MORPHOLOGY

4.1 Word formation

The basis of Sanskrit morphology is the *root*, a morpheme bearing lexical meaning. Through the vowel-gradation processes described above and through the addition of affixes, verbal and nominal *stems* are derived from this root. The grammatical and syntactic identity of a stem in context is then fixed by the addition of an *ending*. In other words, the three major formal elements of the morphology are (i) root, (ii) affix, and (iii) ending; and they are roughly responsible for (i) lexical meaning, (ii) derivation, and (iii) inflection respectively. A (noncompound) word ordinarily contains only one root and one ending, but may have a theoretically unlimited number of affixes. Both ending and affix may also be represented as zero. The canonical structure of a Sanskrit word is thus:

(11) Root – Affix $0-n$ – Ending $0-1$

Numerous examples of roots and their alternants were given above. There are some phonological constraints on root structure, the most important being that no root can end in short *a*, though affixes and endings commonly do, and all roots are monosyllabic (not counting the *i* of *set* roots; see §3.7). There are also some restrictions on co-occurrence of consonants: for example, roots do not contain two aspirates (the historical result of Grassmann's Law) or stops from the same positional series in onset and coda.

Affixes are almost entirely suffixes. There is one infix, alternating *-na/n-* found in a single verbal present class, and one clear prefix, the so-called *augment*, an *a-* prefixed to past tense verb forms in the imperfect and aorist tenses. In addition, the class of *preverbs* mimic prefixes, because they precede a verb (and its nominal derivatives) and modify it semantically (e.g., *ud* "up," *pra* "forth"). In the earliest language, however, the status of these elements is not clear, or rather it fluctuates, as both their position and their accentuation show. In the *Rig-veda* preverbs regularly occur in *tnesis*, in other words, separated from the finite verb. Even when immediately preceding the verb, they maintain their own accent, except in subordinate clauses. This last context is the only one in which they clearly form a part of the phonological word of the finite verb. Preverbs always precede the one undeniable prefix, the *augment*. With nonfinite forms of the verb and with nominal derivatives thereof, preverbs show much clearer univerbation in Vedic, both by position and by accent, and by Classical Sanskrit *tnesis* is no longer possible even with finite forms.

In nominal morphology three elements, *a(n)-* "un," *su-* "well," and *dus-* "ill," function like prefixes, though technically forming compounds, both determinative and possessive.

Besides these few exceptions, suffixes are the rule in affixation. Though there are few absolute phonological constraints on suffixes, most are monosyllabic (though sometimes with the old laryngeal *i* attached, see above) and have relatively simple structure: CV is a common shape. The same is true of endings.

Reduplication is a common morphological process in the verbal system. Although the details cannot be examined, several of the phonological alternation processes discussed above are exemplified in reduplication: dissimilation of aspirates ($\sqrt{dhā}$: *da-dhā-*), alternation of palatals and velars (\sqrt{kr} : *ca-kr-*).

Some words do not conform to the canonical structure. A few forms lack both inflection and root and do not ordinarily serve as derivational bases: for example, the negatives *nā* and *mā*, particles of various functions like *sū* and *hī*, and conjunctions like *ca* and *vā* (some are tonic, some not). Preverbs can be classified here at least originally.

Moreover, a much larger number of words are inflected (and can enter into derivation) but lack a recognizable root. These include many terms of basic vocabulary – kinship terms (e.g., *mātar-* “mother”), body parts (e.g., *nas-* “nose”), flora and fauna (e.g., *śvan-* “dog”) – but are not limited to such semantic categories. Pronouns might be usefully classified here. Numerals also lack roots; some are inflected, some not.

Sanskrit morphology is conveniently divided into two fundamental categories, namely nominal forms and verbal forms, formally distinguished by the types of endings they take and the grammatical categories these endings mark. Adjectives and participles derived from verbs are not formally distinct from nouns; pronouns share the same grammatical categories with nouns, though they may deviate somewhat in inflection. “Adverbs” are usually frozen case forms of adjectives, and nonfinite verbal forms such as infinitives and gerunds also clearly show frozen nominal case endings.

Before discussing nominal and verbal forms separately, we should note certain features and processes they share. Perhaps the most important is the distinction in each between *thematic* and *athematic* inflection. Any *stem*, nominal or verbal, that ends in short *a* (i.e., ends with a suffix consisting of or containing short *a* as final vowel) is thematic. All thematic stems show fixed form throughout their inflection, modified only by the addition of endings. There are no stem alternants and there is no accent shift in the paradigm. Any stem not ending in short *a* is athematic and ordinarily will show stem alternants (as generated by the vowel strengthening patterns discussed above) and often movable accent. For example, the noun stem *deva-* “god” is thematic and maintains this form throughout, whereas *rājan-* “king” is athematic, with the following stem alternants: “strong” *rājān-* (*/rājā-*), “middle” *rājan-*, “weak” *rājñ-* (*/rāja-*). Similarly in verbs, a nonalternating thematic present stem like *bhava-* “become” contrasts with athematic *kṛñó-* / *kṛñu-* (with accent shifted to the ending) “make.” Given the relative simplicity of the former and the frequent morphophonemic complications of the latter, thematic inflection spreads at the expense of athematic inflection during the history of Sanskrit.

Two of the facts noted above – that affixes can be athematic (and alternating) as well as thematic, and that Sanskrit words can contain more than one affix – interact with each other. With very rare exceptions, only one element in any Sanskrit word will alternate within a single paradigm; all the rest will remain frozen in a nonalternating, usually weak form. Whenever a suffix (thematic or athematic) is added to a stem, all preceding elements become frozen. For example, the root $\sqrt{kṛ}$ alternates within its root aorist paradigm: *ákar-am* “I have made” versus *ákr-an* “they have made.” However, when the present-stem alternating suffix *-nó/nu-* is added, the root syllable *kṛ* is fixed in zero-grade: *kṛ-ñó-/kṛ-ñu-*. In turn, with the optative suffix *-yá/-i-* added to that, the present stem is frozen in weak form: *kṛñu-yá-/kṛñv-i-*.

4.2 Nominal morphology

The grammatical categories of Sanskrit nominal forms are gender, number, and case.

4.2.1 Gender

Three genders exist: masculine, neuter, and feminine. Nouns have inherent gender; personal pronouns have no gender, though demonstrative and anaphoric pronouns do. The formal expression is not parallel among the three genders. The feminine is primarily expressed by *derivation*: there are two important feminine-forming suffixes, *-ā-* and *-ī-*. By contrast, the difference between masculine and neuter is primarily *inflectional*. For the most part the

same suffixes form both masculine and neuter nouns, and different case endings signal the different genders. Most stems formally encode masculine versus neuter only in nominative and accusative. A few stem-types (especially *i*-stems) form feminines as well as masculines and neuters, where the feminine is distinguished by different case endings and by the form of modifying adjectives.

4.2.2 Number

Three numbers occur: singular, dual, and plural. The dual is a fully functioning category, used not merely for naturally paired objects, like eyes, but for any collection of two. Notable in Vedic is the "elliptical" dual, with a noun in the dual signalling a conventional paired opposition: for example, *dyāvā*, literally "the two heavens," for "heaven and earth"; *mātarā*, literally "the two mothers," for "mother and father." Number is entirely an inflectional category, except in the personal pronoun.

4.2.3 Case

Sanskrit has eight cases: nominative, accusative, instrumental, dative, ablative, genitive, locative, vocative, though no stems make all eight distinctions in all three numbers. In all stems the dual shows only three distinctions: (i) nominative, accusative, and vocative merge; as do (ii) instrumental, dative, ablative; and (iii) genitive, locative. In all nominal stems the plural collapses nominative and vocative, as well as dative and ablative; only the personal pronouns distinguish dative and ablative in the plural. Even in the singular most stems conflate ablative and genitive; only one nominal stem-type (though the most common, the short *a*-stem) and the pronouns distinguish ablative and genitive singular. Thus, since pronouns lack vocatives, only one stem-type (*a*-stem) has eight distinct case forms in any number. Case function is discussed in §5.

Case is marked inflectionally, by endings, and by stem-form alternations. In alternating paradigms some cases regularly pattern together, in other words, show the same stem alternants. Normally (i) nominative/accusative singular, (ii) nominative/vocative plural and (iii) nominative/accusative/vocative dual (the so-called *strong* cases) operate in opposition to the other, *weak cases* (the terms *direct* versus *oblique* have almost the same range of reference, but are syntactic not formal designations; moreover, the accusative plural is also a direct case).

4.2.4 Nominal stem-classes

Unlike a language such as Latin or Greek, Sanskrit has no closed set of conventionally denoted *noun declensions*. Instead, there is a fairly large set of stem-types, some of which share features of patterning, as well as a sizable group of exceptional stems (not treated here). The first major division is between *root nouns* and *derived nouns*. As the name implies, root nouns combine the bare root, without suffixes, with endings, while derived nouns interpose suffix(es) between root and ending.

4.2.4.1 Vowel stems

The major division in derived nouns is between vowel stems and consonant stems, distinguished by the patterning of stem alternants and to some extent by endings. Among vowel stems we can differentiate three types:

1. The short *a* thematic type, the commonest stem-type in the language, forming masculines (e.g., *deva*- “god”) and neuters (e.g., *phala*- “fruit”). Besides its invariant stem, it is distinguished by somewhat aberrant endings and by the fact that it alone has eight distinct forms in the singular.
2. The *ā* and *ī* feminine stems (e.g., *senā*- “army,” *devī*- “goddess”). In addition to their gender, these stems share a distinctive set of endings in the singular oblique cases.
3. The stems in short *i* and *u*, forming nouns of all three genders (e.g., masc. *agni*- “fire,” fem. *mati*- “thought,” neut. *vāri*- “water”; masc. *paśu*- “cow,” fem. *dhenu*- “milk-cow,” neut. *vasu*- “wealth”). In early Vedic the inflection of all three genders is essentially the same (save for the neuter endings of the direct cases), with weak forms of the stem in the singular direct cases (*agni*-) and strong forms in the singular oblique (*agnay*-). Gradually all three genders develop separate singular oblique forms. The feminine stems become more like the stems of type 2.

4.2.4.2 Consonant stems

A number of varieties occur (*an*-, *ar*-, *ant*-, *vas*-, and *as*-stems, among others), forming primarily masculine and neuter nouns. Most consonant stems share a general patterning tendency: strong forms of the stem occur in the “strong” cases, weak in the “weak” cases (e.g., *rājān*- vs. *rājñ*-; *kartār*- vs. *kartr*-; *sánt*- vs. *sat*-, etc.), in direct opposition to the patterning of the short vowel stems just discussed. A few stem-types show no significant stem alternation (*in*-stems, neuter *s*-stems). Note also that *ar*-stems are often classified as vowel stems (i.e., as *r*-stems), and several of their cases have indeed adopted vowel-stem forms (especially acc. pl., gen. pl.). But the patterning of their stem alternants clearly classifies them with consonant stems, especially *an*-stems.

4.2.5 Endings

Though no scheme of endings is applicable to all stems and all periods of the language, the following chart gives the most common patterns. When there are significant differences, both consonant and vowel-stem endings are given, as well as some feminine alternants.

(12)	Singular				Dual		Plural		
	Cons.	Vow.	Fem.	Neut.	Cons.	Neut.	Cons.	Vow.	Neut.
Nom.	∅	-s		∅	-au	-ī	-as		-Vni
Acc.	-am	-m		[=nom.]	[=nom.]		-as	-Vn	[=nom.]
Instr.	-ā	-nā	-ā		-bhyām			-bhis	
Dat.	-ē	-ē	-āi		[=instr.]			-bhyas	
Abl.	-as	-s	-ās		[=instr.]			[=dat.]	
Gen.		[=abl.]			-os		-ām	-Vnām	
Loc.	-i	var.	-ām		[=gen.]			-su	
Voc.	∅	var.	var.		[=nom.]			[=nom.]	

4.2.6 Comparison of adjectives

There are two different patterns for producing comparatives and superlatives, one primary, that is, by direct attachment to the root, not to a derived adjective (comp. -*īyas*-, splv. -*iṣṭha*-); the other secondary, by attachment to an existing adjective (-*tara*-, -*tama*-). An example of each follows:

- (13) *primary* urú- "wide" váriyas- "wider" váriṣṭha- "widest"
secondary priyá- "dear" priyátara- "dearer" priyátama- "dearest"

In Vedic the secondary suffixes are used rather freely, for example, in compounds like *somapātama-* "most soma-drinking" (i.e., "best drinker of soma"); *ṛtrahantama-* "most *Ṛtra*-smashing" (i.e., "best smasher of *Ṛtra*").

4.2.7 Pronouns

The major division within this category is between (i) the personal pronouns of the first and second persons, unmarked for gender, and (ii) a larger number of gender-distinguishing demonstrative/deictic/anaphoric pronouns and adjectives.

4.2.7.1 Personal pronouns

The cases of these pronouns were noted above, as was the occurrence of a different stem in each number. The number of stems is in fact still greater, in that the first singular and plural and the second plural use a different form for the nominative than for the rest of the paradigm:

(14)	1st sg.	1st pl.	2nd pl.
Nom.	ahám	vayám	yūyám
Elsewhere	m-	asm-	yuṣm-

The other stem formants are 1st dual *āv-*, 2nd. sg. *tu-*, 2nd dual *yuv-*. There also exist enclitic oblique forms, often with yet a different stem (e.g., 1st. pl. *nas*, 2nd pl. *vas*). The endings of the personal pronouns are in part unique to them.

4.2.7.2 Gender-marking pronouns

Such pronouns are characterized by a number of different paradigms and partial paradigms, with different functions sometimes changing over time. Most can be used both as pronouns proper and as demonstrative adjectives. We will mention only the most important and widespread stems, beginning with the strong deictics, nearer *ayám* "this here," farther *asáu* "that yonder." Both have rather aberrant inflection, with an assortment of stems collected from different sources.

The most common pronominal stem is *sá/tám*, with a wide range of uses. While it serves as the anaphoric pronominal par excellence, it also shows traces in early Vedic of deictic usage. Moreover, it is the closest element Sanskrit possesses to both a third-person pronoun and to a definite article. It is also sometimes used with both second- and first-person reference. Its inflection shows archaic inherited features, with initial *s-* in nominative singular (masc. *sá* and fem. *sā*), versus *t-* elsewhere (replicated by Greek masc. *ho*, fem. *hē* [with *h-* < **s-*] but neut. *tó*; see Ch. 24, §4.1.3.4), and with an endingless nominative singular masculine (under certain sandhi conditions).

This stem also shows some peculiarities of inflection, some of which are found also in the stems of the interrogative (*ká-*), the relative (*yá-*), and a class of "pronominal adjectives" such as "other" (*anyá-*), "all" (*víśva-*, replaced by *sárva-*), "one/some" (*éka-*).

4.3 Verbal morphology

Like nouns, verbs are either thematic or athematic. Athematic verbs regularly alternate strong (*guṇa*) forms in the active singular, weak in the rest of the inflection.

The grammatical categories of finite verbs are person, number, voice, tense/aspect, and mood. In general, person/number/voice are expressed by a portmanteau morpheme, the ending; tense/aspect by suffixes, morphological processes directly affecting the root, and/or endings; and mood by suffixes (or endings) following the tense/aspect markers. The canonical shape of a verb is thus:

(15) Root – (Tense/Aspect suffix) – (Mood suffix) – Per./Num./Voice ending

4.3.1 Person and number

These categories index the subject of the verb. There are three persons, first, second, and third (in Western grammatical terminology); and three numbers, singular, dual, and plural. As in the noun, the dual is fully functioning, not limited to subjects naturally occurring in pairs. The nine-member grid defined by these two parameters is the basic building block of the Sanskrit verbal system, the paradigm. Each person/number pair is marked by a separate ending.

4.3.2 Voice

The approach to this topic will differ depending on whether formal or functional aspects are emphasized. Formally, many Sanskrit nine-member paradigms come in matched pairs, in two different voices – with identical stems but different endings. The two voices are *active* and *middle* (or *mediopassive*), or, in the more perspicuous Sanskrit terms, *parasmaipada* “word for another” and *ātmanepada* “word for oneself.” A typical formal configuration, the endings of the present, active, and middle, is given below:

(16)	Active			Middle		
	Singular	Dual	Plural	Singular	Dual	Plural
1st	-mi	-vas	-mas	-e	-vahe	-mahe
2nd	-si	-thas	-tha	-se	-āthe	-dhve
3rd	-ti	-tas	-anti	-te	-āte	-ante

The function of the separate voices is harder to define. Though there exist contrasting pairs such as act. *yajati* “sacrifices (on another’s behalf)” : mid. *yajate* “sacrifices (for one’s own benefit),” which illustrate the Sanskrit terminology, there are other active : middle functional relations: for example, transitive : intransitive, act. *vardhati* “increases X” : mid. *vardhate* “X increases.” Some middles are simply passive in value, though lacking overt passive suffix, and an even greater number have no obvious functional correlate: for example, the numerous *deponents* (to use the Latin term) inflected only in the middle (e.g., *āste* “sits”). The distinction between active and middle is, in the main, a purely formal one synchronically; not surprisingly, the distinction becomes attenuated in the development of the language.

There is, however, an important functional distinction in voice, with various formal encodings: that between active and passive. As just noted, the formal middle sometimes functions as a passive. One particular present-stem type, the suffix-accented *-yá*-present with middle endings, also becomes specialized as a passive (e.g., *ucyate* “is spoken”); and the aorist system contains a third singular of peculiar formation (heavy root syllable and mysterious ending *-i*; type *avāci* “was spoken”), the so-called aorist passive. Passive value is also expressed by several verbal adjectives, the gerundive (“future passive participle”) in *-ya-* and *-tavya-*, and especially the past passive participle in *-ta-* (*-na-*). The latter often substitutes for a finite verb as sentential predicate.

4.3.3 Tense-aspect

The backbone of the tense-aspect system is the three-way contrast between the *present* system, the *aorist* system, and the *perfect* system. Each of these stems produces one or more tenses, as well as (in the early language) moods and participles. The present system has two tenses, the present and the imperfect. In post-Rig-vedic Sanskrit both the aorist and the perfect have only one, though in the *Rig-veda* there is a marginal pluperfect beside the perfect. All three systems can be inflected in either voice:

(17) Stem	Tense
present	present
	imperfect
aorist	aorist
perfect	perfect
	(pluperfect)

Like voice, the tense-aspect system is an elaborate formal edifice whose functional motivations have essentially broken down. Though the system inherited from Proto-Indo-European was an aspectual one, aspect is no longer a clear category even in early Vedic, and only relics of the inherited system can be discerned in the *Rig-veda*. From the Sanskrit point of view, the salient functional distinction is *tense*: present (expressed by the present tense) versus past (expressed by three competing preterital forms, imperfect, aorist, and perfect, as well as by certain nonfinite forms used predicatively).

The old perfect was a stative present functionally; a few Vedic perfects maintain this function, but most already express simple past. The original distinction between the present and aorist systems was probably durative versus punctual, but this can no longer be discerned. Insofar as the aorist can be distinguished from the imperfect in Sanskrit, it expresses immediate past time. The loss of functional distinction among the three past tenses set the stage for the loss of those formal categories in later Indo-Aryan.

4.3.4 Perfect stem morphology

Formally, the perfect is characterized by special endings and, except for one widespread old form (*veda* “knows”), by reduplication. It is built directly to the root, without affixes, and shows ordinary strong/weak stem alternation (type *cakār-a/cakr-úr*). There is only one type of perfect stem formation (except for the “periphrastic perfect” of derivative presents; see §4.3.6).

4.3.5 Primary and secondary endings

The formal distinction between present and aorist systems is less well marked. The endings of the imperfect tense and the aorist are identical (the so-called *secondary* endings), and the endings of the present tense (the *primary* endings) closely resemble these. Compare, for example, the primary and secondary endings of the active singular, and contrast them with the corresponding perfect endings:

(18)	Primary	Secondary	Perfect
1st	-mi	-m	-a
2nd	-si	-s	-tha
3rd	-ti	-t	-a

Unlike the perfect both imperfect and aorist prefix the augment, regularly in Classical Sanskrit and optionally (but commonly) in Vedic. Moreover, several types of stem formation are common to both present and aorist.

4.3.6 Present stem morphology

The indigenous grammarians distinguish ten present classes, which can be conveniently divided into thematic and athematic types. Four thematic classes occur, with the following suffixes added to the root: -a- (Class I); -ā- (VI); -ya- (IV); and -āya- (X). The six athematic classes are as follows: simple root presents (endings added directly to the alternating root, Class II); reduplicated presents (III); and four classes continuing (directly or indirectly) nasal affixes – nasal infix (VII), and suffixed -nó/nu- (V), -ó/u- (VIII), and -ná/nī- (IX). Examples of each follow; thematic forms (with nonalternating stems) are given in the third singular active present, athematic forms in both third singular and third plural active, to display both stem alternants:

(19) Sanskrit present tense classes

I	simple thematic	√bhū “become”	bháva-ti
II	root	√as “be”	ás-ti, s-ánti
III	reduplicated	√hu “pour”	juhó-ti, juhv-ánti
IV	-ya-	√paś “see”	páśya-ti
V	-nó/nu-	√su “press”	sunó-ti, sunv-ánti
VI	-ā-	√viś “enter”	viśá-ti
VII	nasal-infix	√yuj “yoke”	yunák-ti, yuñj-ánti
VIII	-ó/u-	√tan “spread”	tanó-ti, tanv-ánti
IX	-ná/nī-	√krī “buy”	krīṇá-ti, krīṇ-ánti
X	-āya-	√cint “think”	cintāya-ti

There is no longer any clear distinction in function among these various present classes, though again traces of prehistoric distinctions can occasionally be discerned.

Besides the above ten classes, several other formations are formally presents, but are classified separately because they have clear functional correlates.

The *future* is formed with the thematic suffix -syá- (or -īsyá- originally proper to set roots) (e.g., *kariṣyáti* “will do”: √kr̥). There is also a periphrastic future, formed from a noun stem with the -tar- agent suffix.

The so-called secondary conjugations:

1. *Passive*, formed with accented -yá- and middle endings, for example, *nīyáte* “is led”: √nī “lead.” In Classical Sanskrit with the loss of accent the passive cannot be formally distinguished from a middle Class IV present.
2. *Intensive*, formed with heavy reduplication (sometimes disyllabic) and, in later Sanskrit, a -yá- suffix with middle endings. The intensive expresses repeated or intensively performed action, for example, *mármarij-*, *marmrijyáte* “wipe repeatedly, groom”: √mr̥j “wipe.”

3. *Desiderative*, formed with reduplication in *-i-* and a *-sa-* suffix. The desiderative expresses action desired, intended, or about to take place, for example, *pípāsati* “desires, intends, is about to drink”: $\sqrt{pā}$ “drink.”
4. *Causative*, formed with a heavy root syllable and a suffix *-āya-*. Formally not distinguishable from Class X presents, except sometimes in the shape of the root syllable. In the earlier language the causative is ordinarily formed only to intransitive verbs, for example, *pādāyati* “cause to fall”: \sqrt{pad} “fall.”

In addition to present stems built to verbal roots, nouns and adjectives can form *denominative* presents by the addition of the suffix *-yā-*, for example, *áśva-* “horse”: *áśvayāti* “seek horses.”

The above derivative present stems can form a secondary periphrastic perfect, with a feminine accusative singular generated to the present stem, plus the perfect of $\sqrt{kṛ}$ (in the earlier language), $\sqrt{bhū}$ or \sqrt{as} (in the later language), of the type *pādāyām cakāralāsa* “caused to fall.” The periphrastic perfect is especially common with causatives.

4.3.7 Aorist stem morphology

The aorist shares certain stem-types with the present system. The root aorist (e.g., *ābhūt*: $\sqrt{bhū}$ “become”) and thematic aorist (*āvidat*: \sqrt{vid} “find”) resemble Class II and VI presents. Class III presents somewhat resemble the reduplicated aorist, though the aorist has certain formal characteristics (heavy *i-* reduplication, thematic vowel) and a functional connection with the causative (type *āpīpadat* “caused to fall,” parallel to *pādāyati* “causes to fall”) that set it apart.

Proper to the aorist, however, are a variety of sigmatic formations. The *s*-aorist and *iṣ*-aorist were originally identically built, with *s*-suffix, to *aniṣ* and *seṣ* roots respectively. Especially notable in these formations is the consistent *vṛddhi* of the root in the entire active voice, an unusual distribution of grades (e.g., *s*-aor. *ājai-s-* “he conquered”: \sqrt{ji} “conquer”; *āpāvi-ṣ-* “purified”: $\sqrt{pū}$ “purify”). Analogic extensions of these two aorist types led to the creation of the marginal types, *siṣ*-aorist and *sa*-aorist.

The passive aorist was noted in §4.3.2.

4.3.8 Mood

There are four clear moods in early Sanskrit: indicative, imperative, optative, and subjunctive. In addition, the so-called injunctive of early Vedic is considered a mood by some, and the precative, a subtype of the optative, develops in the course of Vedic. This system is reduced by Classical Sanskrit. One global change is the virtual restriction of nonindicative moods to the present stem; in Vedic, aorists and perfects displayed broader modality. Furthermore, the subjunctive is effectively lost, and the injunctive, insofar as it is a mood, becomes restricted in usage.

4.3.8.1 Indicative

The indicative is the unmarked mood, used for statements, questions, etc.

4.3.8.2 Imperative

The imperative expresses command and is marked by special endings on the appropriate tense stem. In Vedic the imperative has a defective paradigm, being found only in second and third persons, but as the subjunctive is lost as a functional category, its first-person

forms are incorporated into the imperative. The negative imperative (i.e., prohibitive) is expressed not by the formal imperative mood, but by the injunctive with a special form of the negative, namely *mā* (not *nā*).

There is also a rare second imperative formation, the so-called *future imperative*, made by adding *-tāt* to the tense stem, expressing a command to be executed after the action of an intervening verb. Its value is usually second singular.

4.3.8.3 Optative

The optative expresses possibility ("might," "could"), necessity ("should," "ought to"), and will/desire ("would"), and is marked by a suffix added to the tense stem. For athematic stems, the suffix is *-yā-* in the active, *-ī-* in the middle, added to the weak stem form (e.g., *s-yā-* to root pres. *as-ti*, *s-anti*: \sqrt{as} ; *kṛṇu-yā-*, *kṛṇv-ī-* to *kṛṇóti*: $\sqrt{kṛ}$). For thematic stems, *-ē-* is substituted for the thematic vowel *-a-* throughout (e.g., *bhāvē-* to thematic pres. *bháva-√bhū*). Both suffixes take secondary endings, with some special details.

The precative is a supercharged optative, primarily expressing desire. It is formed by interposing an *-s-* between the optative suffix and the ending. Thus, the ordinary athematic optative first singular ends in *-yām*, that of the precative in *-yāsam*; that of the first plural optative in *-yāma*, the precative in *-yāsma*.

4.3.8.4 Subjunctive

This mood has disappeared (except for its formal representatives in the imperative) by Classical Sanskrit. It is formed by adding a suffix *-a-* (identical to the thematic vowel) to the tense stem; in thematic verbs this produces a contracted suffix *-ā-* (e.g., *bhāvā-* to *bháva-ti*). Athematic verbs add the *-a-* to their strong forms (e.g., *ás-a-* to *ás-ti*; *kṛṇāv-a-* to *kṛṇó-ti*). The subjunctive stem can take either primary or secondary endings (*āsati*, *ásat*, etc.); in addition, the typical final vowel of primary middle endings, *-e*, is usually strengthened to *-āi* after the Rig-vedic period.

The function of the subjunctive is difficult to define. It often seems to express the future, or volitional future, rather than the more strictly *modal* value its Western name implies. This interpretation fits well with the fact that the future tense is quite rare in early Vedic in finite forms; their place seems to be filled by the subjunctive.

4.3.8.5 Injunctive

Formally the term *injunctive* simply refers to unaugmented preterite forms (i.e., imperfects and aorists). Such forms are quite common in the *Rig-veda* in a variety of contexts, but only one usage persists into later Vedic and Classical Sanskrit: the conjoining of aorist injunctive and the particle *mā* to express prohibitions. Despite the best efforts of numerous distinguished scholars, a common functional core cannot be discerned in the other Rig-vedic contexts, and it seems best to regard these forms as not belonging to a unified modal category, but rather representing a period when the prefixation of the augment was still optional in the preterite.

4.3.9 Nonfinite verbals

Sanskrit possesses a large number of verbal nouns and verbal adjectives, of common occurrence. These ordinarily show verbal syntax (objects in the accusative, for example), and many can stand as the main verb in a clause. Some are built directly to the root, some to tense stems.

4.3.9.1 Infinitive

Classical Sanskrit has a single infinitive, built with the suffix *-tum* added directly to the root in guṇa form (type *kar-tum*: √*kṛ*), which is much rarer in textual usage than the infinitives of other early Indo-European languages. It continues the frozen accusative singular of a nominal stem with a *tu*-suffix, and indeed in Vedic other case forms of this stem appear in infinitival usage: dative *-tave* (/ *-tavāi*), ablative-genitive *-tos*. In addition, other stem-types form infinitives or quasi-infinitives in Vedic, for example, datives to *as*-stems in *-ase*. The line between an infinitive and a simple noun can be difficult to draw in the early language.

Infinitives appear as complements to verbs such as √*śak* “be able” and are used to express purpose. They are neutral as to voice and can express either active (“to X”) or passive (“to be Xed”) value, usually depending on the voice of the form to which they are complement.

4.3.9.2 Gerund

These frozen instrumentals, common in Sanskrit of all periods, are used to express an action prior to (or just simultaneous with) that of the main verb. Standard Classical Sanskrit has two formations, formally distributed: *-tvā* (also made to the *tu*-stem noted under the infinitive, §4.3.9.1) built to an uncompounded root; and *-(t)ya* built to preverb + root (thus the type *kṛ-tvā* vs. *pra-kṛ-tya*). This formal distribution is not always adhered to in the earlier language, and several other related suffixes are also employed.

4.3.9.3 Tense-stem participles

As with the moods, participles tend to become restricted to the present stem in later Sanskrit, although Vedic allows participles to be built to all three tense-aspect stems. Tense-stem participles distinguish voice. The active participle suffix for present and aorist is *-ant-*; the middle suffix for all three tense-aspect stems is *-āna-* for athematic verbs, *-māna-* for thematic. The active perfect participle is made with the suffix *-vas-*, of curious inflection. Though most nonpresent participles disappear by Classical Sanskrit, the perfect participle for *véda* “knows,” *vid-vās-*, survives as an adjective meaning “knowing, wise.”

4.3.9.4 Past passive participle

This is an extremely common form, both as an attributive adjective and as a predicative verb substitute. It is built directly to the unstrengthened root with the suffixes *-tā-*, *-itā-* (originating in *seṭ* roots and still largely found there), *-ná-*, and, rarely, *-vá-*: types *kṛ-tā* “made, done”: √*kṛ* “make, do”; *muṣitā* “stolen”: √*muṣ(i)* “steal”; *san-ná* “seated”: √*sad* “sit”; *pakvā* “cooked, ripe”: √*pac* “cook.” Competing with the three finite past tense forms discussed above, the past passive participle is often the successful contestant, and is responsible for the preterites in a number of later Indo-Aryan languages.

4.3.9.5 Past active participle

Derived from the past passive participle by the addition of the possessive suffix *-vant-* (type *kṛtāvanta-* to *kṛtā-*), it is far less successful than its base.

4.3.9.6 Gerundive (or future passive participle)

The gerundive is another form with passive value, but with the additional component of obligation or necessity (“to be X-ed”), often the equivalent of a passive optative (type *kartavya-* “to be done”). It is formed directly to the root by the addition of one of several suffixes, the most common being *-tavya-* and *-ya-*.

4.4 Compounds

Sanskrit has an extremely well-developed system of nominal compounding; verbal compounding hardly exists. In Vedic, though all types of nominal compounds occur and are frequently encountered, individual compounds are usually limited to two or three members. In Classical Sanskrit, compounds of dozens of members are not infrequent, especially in philosophical texts: the compounding process comes to take the place of the independent syntactic arrangement of inflected words.

4.4.1 Verbal compounds

The verb shows two types of quasi-compounding: (i) the gradual incorporation of preverbs (and functionally equivalent elements) into a verbal complex (type $\sqrt{\text{gam}}$ “go”: $\bar{a}\sqrt{\text{gam}}$ “come”); (ii) the so-called *cvi* construction, which combines nouns and adjectives with both finite and nonfinite forms of the roots $\sqrt{\text{kr}}$ “make” and $\sqrt{\text{bhū}}$ “become” (meaning “make/become X”). In such cases, the nominal first member substitutes invariant *-ī-* for a stem-vowel *-a-* or *-i-*, *-ū-* for *-u-* (e.g., *stambhī-bhavati* “becomes a post”: *stambha-* “post”).

4.4.2 Nominal compounds

Formally, nominal compounding ordinarily involves the concatenation of uninflected words (i.e., stems), resulting in a unit with a single ending and a single accent. The stems may include nouns, adjectives (including participles), adverbs, and pronouns. Both the single ending and the single accent have exceptions in the early language. Inflected case forms may appear in prior compound members, as in *rathe-ṣṭhā-* “standing on a chariot” (with the first member in the locative case). And *paral* compounds (dual dvandvas; see §4.4.2.1) with both members in the dual and both accented (e.g., *mitrā-vāruṇā* “Mitra [and] Varuṇa”) are a well-attested feature of Rig-vedic discourse.

There are three major types of nominal compounds: copulative, determinative, and possessive, known familiarly by their Sanskrit names as *dvandva*, *tatpuruṣa*, and *bahuvrīhi* respectively.

4.4.2.1 Dvandvas

These copulative compounds conjoin two or more stems as parallel members of a series: $X + Y + Z \dots$ (the “lions and tigers and bears” type). Formally the compound may either take the gender of its final member and be inflected as dual or plural (as appropriate), or be treated as a neuter singular collective. In either case the final member is accented (in accented texts). On the Rig-vedic dual dvandvas, with double inflection and double accent, see §4.4.2.

4.4.2.2 Tatpuruṣa

The prior member of this determinative compound limits the following member in some way. Two major subtypes can be distinguished according to the underlying case relations of the members: dependent (*tatpuruṣa* proper) and descriptive (*karmadhāraya*). In the former the prior compound member would be in a different case from that one which follows. A typical relation is genitive + head, as in *nr̥-pāti-*, literally “man-lord,” that is “lord of men”; but other relations are common, especially the limiting of a final past passive participle by an underlying instrumental agent – type *agni-taptā-*, literally “fire-heated,” that is, “heated by fire.” In *karmadhārayas* the prior member is either a qualifier in the same underlying

case as that member which it limits (typically an adjective, i.e., the "black-bird" type) or an adverbial element (*su-* "well," *dus-* "ill," and *a(n)-* "un-" are especially common). The accentual facts of determinatives are complex, but in general the accent falls on the final syllable or the final member.

4.4.2.3 Bahuvrīhi

This possessive compound may be based on any of the preceding types, but adds to the concatenation the semantic feature of possession: the formal sequence $X + Y$ means not simply "X-Y" but "possessing X-Y." English has similar compounds; compare *red-head* and *Bluebeard*.

An important formal consequence of the addition of this semantic feature is that the compound, whose final member is a noun, must be transformed into an adjective, capable of inflection in all genders (hence the common designation "secondary adjective compound"). Sometimes the gender switch can be accomplished silently, as it were, as when neuter nouns in *-a-* simply take masculine endings in the nominative and accusative. Sometimes the adjustment simply requires lengthening or shortening the stem-vowel, as when masculine or neuter nouns in *-a-* become feminized as *ā-* stems or, vice versa, a feminine long *ā-* or *ī-* stem is inflected as a short *a-* or *i-* stem in the masculine or neuter. At other times more complex processes must be employed. These possessive adjectives are then often resubstantivized; bahuvrīhis are a rich source for proper names in Indic and other Indo-European languages (as *Bluebeard* demonstrates).

As with determinatives, the accentual facts are complex, but the accent generally falls on the first member. In accented texts it is thus easy to distinguish determinative compounds from bahuvrīhis, but in later Sanskrit this is not formally possible unless the bahuvrīhi has undergone gender shift.

We might note here that Sanskrit nominal morphology engages in a kind of conspiracy to express the semantic feature "possessing." When a bahuvrīhi cannot be formed, because the notion being expressed is not a compound, a variety of suffixes may be utilized, especially *-vant-* (*-mant-*) and *-in-*, and in early Vedic simple accent shift is possible (e.g., *brāhman-* "formulation" gives *brahmán-* "possessing a formulation," "priest").

4.5 Numerals

The cardinals from 1 to 10, 20, 100, and 1,000 are:

(20)	1	<i>éka-</i>
	2	<i>dvá-</i>
	3	<i>trí-</i>
	4	<i>catúr-</i>
	5	<i>pāñca</i>
	6	<i>ṣaṣ</i>
	7	<i>saptá</i>
	8	<i>aṣṭá</i>
	9	<i>náva</i>
	10	<i>dáśa</i>
	20	<i>viṃśatí</i>
	100	<i>śatá</i>
	1,000	<i>sahásra</i>

The relation of most of these to numerals in other Indo-European languages should be obvious.

There are some unusual inflectional details. *Dvá-* “two” is inflected regularly as a dual in all three genders (masc. nom./acc. *dváu*; fem., neut. *dvé*, etc.). Both *tri-* “three” and *catúr-* “four” display some archaic inflectional features, especially the feminine formant *-sr-* between stem and ending; thus nom./acc. pl. *tisrás* (with dissimilation < **tri-sr-as*), *cátasras*.

Ordinals are derived from cardinals with the suffixes *-ma-* (e.g., *pañcama-* “fifth”) and, rarely, *-tha-* (e.g., *ṣaṣṭha-* “sixth”). Irregular forms include

- | | |
|------------|---|
| (21) first | <i>prathama-</i> |
| second | <i>dvitīya-</i> |
| third | <i>tṛtīya-</i> |
| fourth | <i>turiya-</i> Vedic (< * <i>ktur-</i>), also <i>caturtha-</i> |

5. SYNTAX

Because of its elaborate morphology many traditionally “syntactic” phenomena take place on the level of morphosyntax in Sanskrit. In particular the case system allows the syntactic roles of nominals to be encoded without recourse to rigid word order or obligatory adpositions. Both prepositions and postpositions are rare in early Sanskrit; they become more common later, developing from old preverbs and from frozen case forms of nouns.

5.1 Case usage

Sanskrit cases and their uses are typical of an early Indo-European language: vocative (address); nominative (subject); accusative (direct object; goal of motion; a number of adverbial uses, notably duration of time); instrumental (accompaniment; instrument; agent of the passive; adverbial uses); dative (purpose; indirect object, though the genitive is more commonly used for the latter); ablative (source; cause; comparison); genitive (found in all varieties of adnominal usage; a genitive absolute is also occasionally found, cf. locative absolute); locative (location in both space and time; goal of motion). The locative is also the normal “absolute” case: a noun and modifying participle in the locative can express the time or attendant circumstances under which the action of the main clause occurs: for example, “(on) the sun having risen,” “(on) the enemy fleeing.”

5.2 Word order

Although the case system obviates the need for rigid word order, the order of elements in a Sanskrit sentence is not entirely free. Ordinary prose is SOV (Subject–Object–Verb), with many of the standard typological features of this ordering, such as genitives preceding heads. Poetry and artful prose, however, exploit the opportunities that the syntactic clarity of the morphological system affords, by thoroughly scrambling the order of elements for expressive or discourse purposes. Even in the most extreme examples, however, it is usually possible to formulate principles of movement from a putative underlying order parallel to simple prose.

Overt marking of the subject is not necessary; the bare verb, with person/number markings, is sufficient. First- and second-person subject pronouns are used in addition to the verb only for emphatic or contrastive value. The third-person “pronoun” *sá* is more frequent with third-person verbs, but it ordinarily serves discourse functions: anaphoric to a noun previous in the discourse or coreferent with the relative pronoun in a subordinate clause.

Not only finite verb forms but also participles, especially the past passive participle, can fill the slot V. In this case the copula normally appears only in the first and second persons, and even in those circumstances the personal pronoun can serve instead:

- | | | | | |
|------|--------------------------------|----|----------------------------------|--------------------|
| (22) | <i>gató 'smi</i> (with copula) | or | <i>ahám gatáh</i> (with pronoun) | "I went" |
| | <i>gató 'si</i> (with copula) | or | <i>tvám gatáh</i> (with pronoun) | "you went" |
| | but | | <i>gataḥ</i> | "he [she/it] went" |

Also common are nominal sentences – that is, the predication of a noun (or adjective) to a noun (or pronoun) without an overt copula.

5.3 Cliticization

As in other early Indo-European languages, sentences frequently begin with a chain of clitics attached to the initial, accented word of the sentence, occupying "Wackernagel's position." Such a chain of clitics (and pseudo-clitics – some carry accent) consists of sentential particles (often several to the sentence), conjunctions, and pronouns fronted from their underlying position in the clause; their order is determined by both syntactic class and phonological shape. Word-level conjunctions and pronominal clitics may also appear elsewhere in the clause, the latter ordinarily attached to their head. In such positions the pronoun may either precede or follow the word it is attached to, but clause-initial proclitics are not permitted: all clauses (and their metrical equivalent, verse lines) begin with an accented word. Especially common initial hosts include coordinating and subordinating conjunctions, preverbs in tmesis, and tonic demonstrative and anaphoric pronouns. Much recent work on Sanskrit syntax has concentrated on the constituents of this initial chain and their functions.

5.4 Subordination

A fully inflected relative pronoun *yá-* and a number of subordinating conjunctions built to this stem (*yadá* "when," *yádi* "if," etc.) mark subordinate clauses. These elements are normally fronted (*wh-movement*) from wherever they originate in the clause, but as other elements (including entire constituents) can be topicalized around them, the fronting is sometimes not superficially obvious. Relative clauses either precede or follow the main clause (the former is more usual except in the case of relative clauses of purpose); there is almost no embedding.

In early Vedic, subordinate clauses are sometimes marked only by verbal accentuation, not by a subordinating conjunction; and some particles (notably *hi*) also induce verbal accentuation, presumably a mark of subordination.

Indirect discourse is quite rare, especially in Classical Sanskrit; such clauses are usually expressed by direct discourse marked by the clause-final quotative particle *iti*. For example, "he thought that he would go" would be expressed as "he thought, 'I will go.'"

Other, nonclausal types of subordination are quite common. For example, a series of gerunds with nominal complements is often completed by a single finite verb (type "having come, having asked the king for permission, having received it, he went away"). A notable feature of the syntax of the gerund is that its subject is the logical *agent* of the main clause, not necessarily the overt grammatical subject (type "having smashed [ger.] with a cudgel, the tiger [nom.] was killed by the man [instr.]", where the subject of the gerund is "the man" in the instrumental).

Participles and possessive compounds often correspond to relative clauses in other languages. Noteworthy is the use of the present participle of the verb “to be” (*sánt-*) as a concessive marker (“*although* being X, ...”). Bahuvrīhis often serve as nonrestrictive relative clauses (type “Indra, [lit.] possessing slain Vṛtra”, i.e., “who had slain Vṛtra”).

Unlike some other early Indo-European languages, Sanskrit has no elaborate rules governing the succession of moods and tenses in conditional sentences.

5.5 Agreement

The usual agreement rules of early Indo-European languages hold for Sanskrit: subjects agree with their verbs in person and number; adjectives with the nouns they modify in number, gender, and case; relative pronouns with their antecedents in number and gender.

There are a few interesting exceptions. The well-known Ancient Greek rule, whereby a neuter plural subject takes a singular verb, is preserved only in a few Vedic relics; ordinarily a plural verb is used. Vedic prose has developed a subtype of defining relative clause (type: “... the X, which is Y”) in which the relative marker is always neuter singular *yád*, whatever the gender and number of X and Y. This usage is reminiscent of the Iranian *izafe* marker, which has developed from the same form, but it is not clear if the two constructions are directly related. In some other equational nominal clauses, by contrast, an anaphoric pronoun is attracted to the number and gender of its antecedent.

Though conjoined nominals ordinarily agree in case, an apparently inherited exception in Vedic involves the conjoining of vocatives by *ca* “and,” where the second underlying vocative appears instead in the nominative. This phenomenon is denominated the *vāyav indraś ca* construction after one of its principal examples (“o Vāyu [voc.] and Indra [nom.]”).

5.6 Stylistic syntactic developments

One may consider the history of Sanskrit a history of style, and style in turn is linked to textual genre. Although neither the grammar nor the syntax of Sanskrit shows any significant changes after the fixation of the language by the early grammarians, the usage of these fixed elements significantly alters its balance in the Classical period. The emphasis falls heavily on the nominal system, and the complex verbal system outlined above is exploited far less. We have already noted some of the features of this change in emphasis – the efflorescence of the compounding system, the employment of nominal formations built to verbal roots in preference to finite verbs, the expansion of the adnominal case, the genitive. Sanskrit works of “high” style, court literature and philosophical discourse, take these tendencies to remarkable extremes, while technical treatises, with an eye to verbal economy, arrive at a similar nominal style from a somewhat different angle.

6. LEXICON

A very large proportion of the Sanskrit vocabulary in all periods consists of transparent Indo-European inheritances. Examples need hardly be given; but the numerals given above (§4.5), as well as kinship terms like *pitar* “father,” *māter*, “mother,” *sūnu* “son,” *duhitar* “daughter” can serve as illustrations. Not surprisingly, however, even earliest Vedic has words without clear Indo-European correspondences. While some of these may nonetheless still continue Proto-Indo-European etyma, others doubtless were borrowed from languages with which

the Sanskrit speakers came in contact. The difficulty is determining the source languages, given the fact that we have no records of likely languages from remotely the same era. Though Sanskrit speakers no doubt encountered speakers of Dravidian language(s), no Dravidian language is attested until around the beginning of the present era, and then only in South India. We do not know what a northern Dravidian language would have looked like in the second millennium BC. Our knowledge of the Muṇḍa languages (belonging to the Austro-Asiatic family) comes only from the modern era. Many scholars have proposed Dravidian and Muṇḍa sources for Sanskrit words (and indeed phonemes, syntactic constructions, and so on). It is reasonable to accept the principle, but difficult to judge the plausibility of any particular suggestion. Even when a single etymon clearly reveals itself in Sanskrit and one or more Dravidian languages, for example, borrowing may have gone in the other direction, or both families may have borrowed from a third source. Later (i.e., post-Vedic) Dravidian borrowings into Sanskrit are less controversial.

In addition to borrowing from non-Indo-Aryan languages, Sanskrit also sometimes reincorporates vocabulary showing Middle Indic phonological developments, often with some phonological hypercorrection.

7. READING LIST

The standard synchronic grammar of Sanskrit in English is Whitney 1889, which, along with its supplement, Whitney 1885, is invaluable. The standard historical grammar is the multivolume but still unfinished (lacking the verb) Wackernagel and Debrunner 1896–. The first volume, reissued in 1957, has a detailed general introduction to the language by L. Renou. Many of Renou's other works can be consulted with profit, including his short but elegant history of the language (1956). The classic work on syntax (but only of the Vedic period) is Delbrück 1888. Speijer 1886 treats the Classical language. The standard etymological dictionary is Mayrhofer 1956–1976, currently updated and significantly expanded in Mayrhofer 1986–. A general discussion of the language, though with personal views, is found in Burrow 1955. A short survey, along the same lines as this, is found in Cardona 1987. Both Bloch 1965 and Masica 1991, though concentrating on later Indo-Aryan, nonetheless treat many aspects of Sanskrit as starting points for later developments.

Bibliography

- Bloch, J. 1965. *Indo-Aryan from the Vedas to Modern Times*. Translated and revised by A. Master. Paris: Adrien-Maisonneuve.
- Burrow, T. 1955. *The Sanskrit Language*. London: Faber and Faber.
- Cardona, G. 1987. "Sanskrit" In B. Comrie (ed.), *The World's Major Languages*, pp. 448–469. New York: Oxford University Press.
- Delbrück, B. 1888. *Altindische Syntax*. Reprint 1968, Darmstadt: Wissenschaftliche Buchgesellschaft.
- Masica, C. P. 1991. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
- Mayrhofer, M. 1956–1976. *Kurzgefasstes etymologisches Wörterbuch des Altindischen*. Heidelberg: Carl Winter.
- . 1986–EWA. = *Etymologisches Wörterbuch des Altindischen*. Heidelberg: Carl Winter.
- Renou, L. 1956. *Histoire de la langue sanskrite*. Lyon: Editions IAC.
- . 1957. *Introduction générale, Altindische Grammatik*, vol. I [see Wackernagel 1896].
- Speijer, J. S. 1886. *Sanskrit Syntax*. Leiden: Brill.
- Wackernagel, J. and A. Debrunner, *Altindische Grammatik*. Göttingen: Vandenhoeck and Ruprecht. Vol. I 1896. *Lautehre* (J. Wackernagel). Reprint 1957, with intro. by L. Renou.

-
- Vol. II.1 1905. *Einleitung zur Wortlehre. Nominalkomposition* (J. Wackernagel). Reprint 1957.
Vol. II.2 1954. *Die Nominalsuffixe* (A. Debrunner).
Vol. III 1930. *Nominalflexion – Zahlwort – Pronomen* (A. Debrunner and J. Wackernagel).
Whitney, W. D. 1885. *The Roots, Verb-Forms and Primary Derivatives of the Sanskrit Language*.
Leipzig: Breitkopf and Härtel.
———. 1889. *Sanskrit Grammar* (2nd edition). Cambridge, MA: Harvard University Press.

